

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: March 31, 2004, 13:52:29 ; Search time 3493 Seconds
(without alignments)
7167.840 Million cell updates/sec

Title: US-09-975-856-1
Perfect score: 576
Sequence: 1 ATGACGAGACGACGCTT.....ATGACGAGTAACCTCCCTCG 576

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	576	100.0	576	6	ARI67456	ARI67456 Sequence
2	576	100.0	576	6	ARI70503	ARI70503 Sequence
3	576	100.0	576	6	ARI82977	ARI82977 Sequence
4	576	100.0	576	6	AR305671	AR305671 Sequence
5	576	100.0	576	6	AX740218	AX740218 Sequence
6	576	100.0	576	6	BD096872	BD096872 Isolated
7	576	100.0	576	6	BD218420	BD218420 SSX gene,
8	576	100.0	576	6	HSU90841	U90841 Homo sapien
9	576	100.0	576	6	BC005325	BC005325 Homo sapi
10	574.4	99.7	1250	9	BC005325	A48452 Sequence 17
11	496.6	86.2	766	6	AX821947	AX821947 Sequence
12	496.6	86.2	766	9	HSSSX2	X86175 H.sapiens m
13	496.6	86.2	931	6	AR025465	AR025465 Sequence
14	496.6	86.2	931	6	AR060380	AR060380 Sequence
15	496.6	86.2	931	6	ARI17874	ARI17874 Sequence
16	496.6	86.2	931	6	ARI67453	ARI67453 Sequence
17	496.6	86.2	931	6	I85576	I85576 Sequence 2
18	496.6	86.2	931	6	AR287596	AR287596 Sequence
19	496.6	86.2	931	6	AR305668	AR305668 Sequence
20	496.6	86.2	931	6	AX114023	AX114023 Sequence
21	496.6	86.2	931	6	AX719099	AX719099 Sequence
22	496.6	86.2	931	6	BD218417	BD218417 SSX gene,
23	496.6	86.2	1309	6	AX331558	AX331558 Sequence
24	496.6	86.2	1309	6	AX331950	AX331950 Sequence
25	496.6	86.2	1309	9	HSXD21LEK	Z49105 H.sapiens H
26	496	86.1	1249	9	HSU90840	U90840 Homo sapien
27	495	85.9	1276	9	BC016957	BC016957 Homo sapi
28	493.4	85.7	1282	9	BC007343	BC007343 Homo sapi
29	492.8	85.6	1282	9	S82471	S82471 Homo sapien
30	482.2	83.7	576	6	ARI67457	ARI67457 Sequence
31	482.2	83.7	576	6	ARI70504	ARI70504 Sequence
32	482.2	83.7	576	6	ARI82978	ARI82978 Sequence
33	482.2	83.7	576	6	AR305672	AR305672 Sequence
34	482.2	83.7	576	6	AX740219	AX740219 Sequence
35	482.2	83.7	576	6	BD096873	BD096873 Isolated
36	482.2	83.7	576	6	BD218421	BD218421 SSX gene,
37	481.6	83.6	1261	9	BC001003	BC001003 Homo sapi
38	481.2	83.5	766	6	A48450	A48450 Sequence 15
39	480.6	83.4	576	9	HSU90842	U90842 Homo sapien
40	480	83.3	766	6	ARI67452	ARI67452 Sequence
41	480	83.3	766	6	AR305667	AR305667 Sequence
42	480	83.3	766	6	AX821946	AX821946 Sequence
43	480	83.3	766	6	BD218416	BD218416 SSX gene,
44	480	83.3	766	9	HSSSX1	X86174 H.sapiens m
45	433.2	75.2	1363	9	BC016640	BC016640 Homo sapi

ALIGNMENTS

RESULT 1
ARI67456
LOCUS
DEFINITION Sequence 5 from patent US 6287756.
ACCESSION ARI67456
VERSION ARI67456.1 GI:17903237
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 576)
Tureci,O., Chen,Y.-T., Sahin,U., Gure,A.O., Old,L.J. and
pfreundschuh,M.
AUTHORS
METHODS for determining presence of cancer in a sample by
determining expression of an SSX gene
PAT 17-DEC-2001

```

Query Match      100.0%; Score 576; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.6e-164;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Ov 1 ATGACCGAGACGACGCGCTTTGCAAGGAGACCCGAGGATGTCGCAATATCAGAGAAG 60

```

QY	121	AAATCTCGGAGAAAATCGTCTATGTGTATATGAAGCTAAACTATGAGGTTCATGACTAAA	180
Db	121	AAATCTCGGAGAAAATCGTCTATGTGTATATGAAGCTAAACTATGAGGTTCATGACTAAA	180
QY	181	CTAGGTTTCAAGGTCAACCTTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC	240
Db	181	CTAGGTTTCAAGGTCAACCTTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC	240
QY	241	GGGAATGATTTTGGTAAACGATCGAAACCCACAGGAATCAGGTTGAAACGTCCTCAGATGACT	300
Db	241	GGGAATGATTTTGGTAAACGATCGAAACCCACAGGAATCAGGTTGAAACGTCCTCAGATGACT	300
QY	301	TTCGGAGCCTCCAGAGAAATCTTCCGAAAGATCATGCCCAAGAACCCACAGAGAGAGAA	360
Db	301	TTCGGAGCCTCCAGAGAAATCTTCCGAAAGATCATGCCCAAGAACCCACAGAGAGAGAA	360
QY	361	AATGGTTTGAAGGAAGTGCACAGAGGCATCTGGGCCACAAAAATGATGGGAAACAGCTGTGC	420
Db	361	AATGGTTTGAAGGAAGTGCACAGAGGCATCTGGGCCACAAAAATGATGGGAAACAGCTGTGC	420
QY	421	CCCCGGGAAATCCAAGTACTCTTGGAGAAAGATTAAAGACATCTGGACCCAAAAGGGGG	480
Db	421	CCCCGGGAAATCCAAGTACTCTTGGAGAAAGATTAAAGACATCTGGACCCAAAAGGGGG	480
QY	481	AAACATCCCTGGACCCACAGACTCCGTGAGAGAAAGCAGCTGGTGGTTTATTAAGAGATC	540
Db	481	AAACATCCCTGGACCCACAGACTCCGTGAGAGAAAGCAGCTGGTGGTTTATTAAGAGATC	540
QY	541	AGCGACCTCGAGGAAGATGACGAGTAATCCCTCCG	576
Db	541	AGCGACCTCGAGGAAGATGACGAGTAATCCCTCCG	576

RESULT 5
 AX740218
 LOCUS
 DEFINITION Sequence 1 from Patent.EP1300463.

576 bp
 mRNA
 linear
 PAT 08-MAY-2003

```

KEYWORDS
SOURCE
  ORGANISM
    Homo sapiens (human)
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1
  Gure, A.O., Tureci, O., Sahin, U., Tsang, S., Scanlan, M.J., Knuth, A.,
  Pfeundscht, M., Old, L.J. and Chen, Y.T.
  Isolated nucleic acid molecules encoding sxx family members and
  uses thereof
  Patent: EP 1300463-A 1 09-APR-2003;
  LUDWIG INSTITUTE FOR CANCER RESEARCH (US) ; MEMORIAL
  SLOAN-KETTERING CANCER CENTER (US) ; CORNELL RESEARCH FOUNDATION
  (US)
FEATURES
  source
    Location/Qualifiers
      1..576
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
      1..576
        /gene="sxx4"
  gene

```

```

ORIGIN
      100      200      300      400      500      600      700      800      900      1000
1   /gene="SSX4"
2   /codon_start=1
3   /protein_id="CAD90570.1"
4   /db_xref="GI:30519359"
5   /db_xref="REMTREMBL:CAD90570"
6   /translation="MNGDAFAARRRDQAQISEKLKAFDDIAKTFYSKKEWEMKMSK
7   KIVVYVYKLYEVMTKLGFKVTLPPFMSKRAADFGHNDFGNDRHNRHNOVERPQPTF
8   SIQELFPMKIMSKPAEENGKVEPEASQPQDNGKQLCPPGNPSTLEKINKTSGPKR
9   KXWTHRLEKRLVYVEISEIDPEEDDE"
      1100      1200      1300      1400      1500      1600      1700      1800      1900      2000

```

Best Local Similarity 100.0%; Pred. No. 2.6e-164;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGATGATCTCAATATCAGAGAAG 60
DB 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGATGATCTCAATATCAGAGAAG 60
QY 61 TTACGAAGGCTTCGATGATATTGCCAATATCTCTTAAGAAAGAGTGGAAAGATG 120
DB 61 TTACGAAGGCTTCGATGATATTGCCAATATCTCTTAAGAAAGAGTGGAAAGATG 120
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAATATGAGTCACTATAA 180
DB 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAATATGAGTCACTATAA 180
QY 181 CTAGTTTCAAGGTCAACCTCCGACCTTCATGCTAGTAAACGGCTGAGACTCCAC 240
DB 181 CTAGTTTCAAGGTCAACCTCCGACCTTCATGCTAGTAAACGGCTGAGACTCCAC 240
QY 241 GGGATGATTTTGGTAAGCATCGAAACACAGGAATCAGGTTGAACCTCTCAGATGACT 300
DB 241 GGGATGATTTTGGTAAGCATCGAAACACAGGAATCAGGTTGAACCTCTCAGATGACT 300
QY 301 TTGGCAGCTCCAGAGAAATCTTCCGAAGATCATGCCAAGAGCCAGCAGAGAGAA 360
DB 301 TTGGCAGCTCCAGAGAAATCTTCCGAAGATCATGCCAAGAGCCAGCAGAGAGAA 360
QY 361 AATGTTTGAAGGAGTCCAGAGCATCTCGGCCACAAATATGATGGAAACAGCTGTC 420
DB 361 AATGTTTGAAGGAGTCCAGAGCATCTCGGCCACAAATATGATGGAAACAGCTGTC 420
QY 421 CCCCAGGAAATCCAAAGTACCTTGGAGAAATTAACAGAGATCTGAGACCATCGAG 480
DB 421 CCCCAGGAAATCCAAAGTACCTTGGAGAAATTAACAGAGATCTGAGACCATCGAG 480
QY 481 AAACATGCTGAGACCCACAGATCGCTGAGAGAAAGAGCTGGTGTATGAAGATC 540
DB 481 AAACATGCTGAGACCCACAGATCGCTGAGAGAAAGAGCTGGTGTATGAAGATC 540
QY 541 AGCGACCTCGAGAAAGATGACGAGTAATCCCTCG 576
DB 541 AGCGACCTCGAGAAAGATGACGAGTAATCCCTCG 576

RESULT 6
BD096872
LOCUS
DEFINITION
Isolated nucleic acid molecules encoding SSX family members and uses thereof.
ACCESSION
BD096872
VERSION
JP 2001527408-A/1.
KEYWORDS
unidentified
SOURCE
unidentified
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 576)
AUTHORS
Gure,A.O., Tureci,O., Sahin,U., Teang,S., Scanlan,M.J., Knuth,A., Pfreundschuh,M., Old,L.J. and Chen,Y.T.
TITLE
Isolated nucleic acid molecules encoding SSX family members and uses thereof
JOURNAL
Patent: JP 2001527408-A 1 25-DEC-2001;
LUDWIG INSTITUTE FOR CANCER RESEARCH, MEMORIAL SLOAN KETTERING
CANCER CENTER, CORNELL RESEARCH FOUNDATION
COMMENT
OS Unidentified
PN JP 2001527408-A/1
PD 25-DEC-2001
PF 25-FEB-1998 JP 1998548050
PR 05-MAY-1997 US 08/851138
PI ALI O GURE, OZLEM TURECI, UGUR SAHIN, SOLAM TSANG, MATTHEW J PI
SCANLAN,
PI ALEXANDER KNUTH, MICHAEL PFEUNDSCHEH, LLOYD J OLD, YAO TSENG PI
CHEN
PC C12N5/10, C12N15/12, C12P21/02, C12Q1/68

CC Strandedness: Single;
CC Topology: linear;
CC Isolated nucleic acid molecules encoding SSX family members
CC thereof
FH Key Location/Qualifiers
FT source 1..576
FT /organism='Unidentified',
FT Location/Qualifiers
FT 1..576
FT /organism='unidentified'
FT /mol_type='genomic DNA'
FT /db_xref='taxon:32644'

FEATURES
source
ORIGIN
Query Match 100.0%; Score 576; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.6e-164;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGATGATCTCAATATCAGAGAAG 60
DB 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGATGATCTCAATATCAGAGAAG 60
QY 61 TTACGAAGGCTTCGATGATATTGCCAATATCTCTTAAGAAAGAGTGGAAAGATG 120
DB 61 TTACGAAGGCTTCGATGATATTGCCAATATCTCTTAAGAAAGAGTGGAAAGATG 120
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAATATGAGTCACTATAA 180
DB 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAATATGAGTCACTATAA 180
QY 181 CTAGTTTCAAGGTCAACCTCCGACCTTCATGCTAGTAAACGGCTGAGACTCCAC 240
DB 181 CTAGTTTCAAGGTCAACCTCCGACCTTCATGCTAGTAAACGGCTGAGACTCCAC 240
QY 241 GGGATGATTTTGGTAAGCATCGAAACACAGGAATCAGGTTGAACCTCTCAGATGACT 300
DB 241 GGGATGATTTTGGTAAGCATCGAAACACAGGAATCAGGTTGAACCTCTCAGATGACT 300
QY 301 TTGGCAGCTCCAGAGAAATCTTCCGAAGATCATGCCAAGAGCCAGCAGAGAGAA 360
DB 301 TTGGCAGCTCCAGAGAAATCTTCCGAAGATCATGCCAAGAGCCAGCAGAGAGAA 360
QY 361 AATGTTTGAAGGAGTCCAGAGCATCTCGGCCACAAATATGATGGAAACAGCTGTC 420
DB 361 AATGTTTGAAGGAGTCCAGAGCATCTCGGCCACAAATATGATGGAAACAGCTGTC 420
QY 421 CCCCAGGAAATCCAAAGTACCTTGGAGAAATTAACAGAGATCTGAGACCATCGAG 480
DB 421 CCCCAGGAAATCCAAAGTACCTTGGAGAAATTAACAGAGATCTGAGACCATCGAG 480
QY 481 AAACATGCTGAGACCCACAGATCGCTGAGAGAAAGAGCTGGTGTATGAAGATC 540
DB 481 AAACATGCTGAGACCCACAGATCGCTGAGAGAAAGAGCTGGTGTATGAAGATC 540
QY 541 AGCGACCTCGAGAAAGATGACGAGTAATCCCTCG 576
DB 541 AGCGACCTCGAGAAAGATGACGAGTAATCCCTCG 576

RESULT 7
BD218420
LOCUS
DEFINITION
SSX gene, method of determining the occurrence of cancer in sample by determining the expression of peptides originating in the SSX gene and NY-ESO-1 gene and utilization thereof.
ACCESSION
BD218420
VERSION
JP 2002519013-A/5.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

BD218420 576 bp DNA linear PAT 17-JUL-2003
SSX gene, method of determining the occurrence of cancer in sample by determining the expression of peptides originating in the SSX gene and NY-ESO-1 gene and utilization thereof.

REFERENCE 1 (bases 1 to 576)
 Tureci, O., Sahin, U., Pfeundschoh, M., Rammensee, G., Stevanovic, S.,
 Chen, Y.T., Gure, A. and Old, L.J.
 SSX gene, method of determining the occurrence of cancer in sample
 by determining the expression of peptides originating in the SSX
 gene and NY-ESO-1 gene and utilization thereof
 Patent: JP 2002519013-A 5 02-JUL-2002;
 LUDWIG INSTITUTE FOR CANCER RESEARCH
 OS Homo sapiens (human)
 PN JP 2002519013-A/5
 PD 02-JUN-2002 JP 2000557145
 PF 25-JUN-1999 JP 2000557145
 PR 26-JUN-1998 US 09/105839
 PI OZLEM TURECI, UGUR SAHIN, MICHAEL PFEUNDSCHUH, GEORG RAMMENSEE,
 PI STEFAN STEVANOVIC, YAO TSENG CHEN, ALI GURE, ILOYD J OLD PC
 C12N15/09, A61K38/00, C07H21/04, C12N5/10, C12P21/04, C12Q1/68, PC
 G01N33/48
 PC G01N33/53, G01N33/574, C12N15/00, A61K37/02, C12N5/00 CC SSX
 gene, method of determining the occurrence of cancer in CC
 sample by
 CC determining the expression of peptides originating in the SSX
 gene and
 CC NY-ESO-1 gene and utilization thereof
 FH Key Location/Qualifiers
 FT source 1..576
 FT /organism='Homo sapiens (human)'
 FT Location/Qualifiers
 FEATURES
 source
 1..576
 /organism='Homo sapiens'
 /mol_type='genomic DNA'
 /db_xref='taxon:9606'
 ORIGIN
 Query Match 100.0%; Score 576; DB 6; Length 576;
 Best Local Similarity 100.0%; Pred. No. 2.6e-164;
 Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ATGAACGAGACGAGCGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
 Db 1 ATGAACGAGACGAGCGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
 Qy 61 TTACGAAAGGCGCTTCGATGATATTGCCAATATCTTCTAAGAAAGAGTGGGAAAGATG 120
 Db 61 TTACGAAAGGCGCTTCGATGATATTGCCAATATCTTCTAAGAAAGAGTGGGAAAGATG 120
 Qy 121 AAATCCTCGGAGAAAACGCTATGTGTATGATGAAGTAACTATGAGGTCATGACTAAA 180
 Db 121 AAATCCTCGGAGAAAACGCTATGTGTATGATGAAGTAACTATGAGGTCATGACTAAA 180
 Qy 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
 Db 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
 Qy 241 GGGAAATGATTTTGGTAACGATCGAAACACAGGAATCAGGTTGAACGCTCCTCAGATGACT 300
 Db 241 GGGAAATGATTTTGGTAACGATCGAAACACAGGAATCAGGTTGAACGCTCCTCAGATGACT 300
 Qy 301 TTGCGCAGCTCCACAGAGATCTTCCGGAAGATCATGCCAAGAGCCAGACAGAGAGAA 360
 Db 301 TTGCGCAGCTCCACAGAGATCTTCCGGAAGATCATGCCAAGAGCCAGACAGAGAGAA 360
 Qy 361 AATGTTTGAAGGAAGTCGAGAGCATCTGGCCACAAAATGATGGGAAACAGCTGTGC 420
 Db 361 AATGTTTGAAGGAAGTCGAGAGCATCTGGCCACAAAATGATGGGAAACAGCTGTGC 420
 Qy 421 CCCCAGGAAATCCAGTACCTTGGAGAGATTAAAGAGATCTGGACCCAAAGGGGG 480
 Db 421 CCCCAGGAAATCCAGTACCTTGGAGAGATTAAAGAGATCTGGACCCAAAGGGGG 480
 Qy 481 AAACATGCTCGGACCCACAGACTGCGTGAGAGAAAAGCAGCTGGTGGTTTATGAAGATC 540
 Db 481 AAACATGCTCGGACCCACAGACTGCGTGAGAGAAAAGCAGCTGGTGGTTTATGAAGATC 540

Qy 541 AGGACCCCTGAGGAGATGACGAGTAACCTCCCTCG 576
 Db 541 AGGACCCCTGAGGAGATGACGAGTAACCTCCCTCG 576
 RESULT 8
 HSU90841
 LOCUS Homo sapiens SSX4 (SSX4) mRNA, complete cds. PRI 18-MAR-1998
 DEFINITION Homo sapiens SSX4 (SSX4) mRNA, complete cds.
 ACCESSION U90841
 VERSION U90841.1 GI:2952022
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 576)
 Gure, A.O., Tureci, O., Sahin, U., Tsang, S., Scanlan, M.J., Jager, E.,
 Knuth, A., Pfeundschoh, M., Old, L.J. and Chen, Y.T.
 SSX: a multigene family with several members transcribed in normal
 testis and human cancer
 Int. J. Cancer 72 (6), 965-971 (1997)
 JOURNAL 98021352
 MEDLINE 9378559
 PUBMED
 REFERENCE 2 (bases 1 to 576)
 Gure, A.O., Tureci, O., Sahin, U., Tsang, S., Scanlan, M., Knuth, A.,
 Pfeundschoh, M., Old, L.J. and Chen, Y.T.
 Direct Submission
 Submitted (05-FEB-1997) Pathology C-320, Cornell University, 1300
 York Ave., New York City, NY 10021, USA
 LOCATION/Qualifiers
 1..576
 /organism='Homo sapiens'
 /mol_type='mRNA'
 /db_xref='taxon:9606'
 1..576
 /gene='SSX4'
 1..567
 /note='putative translation product; Kruppel-associated
 box containing SSX gene'
 /codon_start=1
 /product='SSX4'
 /protein_id='AAC05820.1'
 /db_xref='GI:2952023'
 translation='MNGDDAFARRPRDDAQISEKRAFDDIAKYFSKKEWKKMSSE
 KIVYVMKLYEVMVTKGFVTLPPFMRSKRAADPHGNDFGNDRNHRNQVERPQMTFG
 SIQRIFPKIMPKPAEENGLKEVPEASQPDGKQLCPGPNPSTLEKINKTSGPKRG
 KAWTHRLRRKQLVVEISDPEDDE'
 1..18
 /gene='SSX4'
 /note='corresponds to the 5' PCR primer that was used to
 isolate this gene and is identical to SSX2, GenBank
 Accession Number X86175; therefore, the SSX4 gene sequence
 might contain differences within this span'
 primer_bind
 100.0%; Score 576; DB 9; Length 576;
 Best Local Similarity 100.0%; Pred. No. 2.6e-164;
 Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 ORIGIN
 Query Match 100.0%; Score 576; DB 9; Length 576;
 Best Local Similarity 100.0%; Pred. No. 2.6e-164;
 Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ATGAACGAGACGAGCGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
 Db 1 ATGAACGAGACGAGCGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
 Qy 61 TTACGAAAGGCGCTTCGATGATATTGCCAATATCTTCTTAAAGAGAGTGGGAAAGATG 120
 Db 61 TTACGAAAGGCGCTTCGATGATATTGCCAATATCTTCTTAAAGAGAGTGGGAAAGATG 120
 Qy 121 AAATCCTCGGAGAAAACGCTATGTGTATGATGAAGTAACTATGAGGTCATGACTAAA 180
 Db 121 AAATCCTCGGAGAAAACGCTATGTGTATGATGAAGTAACTATGAGGTCATGACTAAA 180

QY 181 CTAGGTTTCAAGTCACCCCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
 Db 181 CTAGGTTTCAAGTCACCCCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
 QY 241 GGGAAATGATTTTGTATACGATCGAAACACACAGGAATCAGGTTGAAAGTCTTCAGATGACT 300
 Db 241 GGGAAATGATTTTGTATACGATCGAAACACACAGGAATCAGGTTGAAAGTCTTCAGATGACT 300
 QY 301 TTCCGGCAGCCTCCAGAGAAATCTTCCCGAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360
 Db 301 TTCCGGCAGCCTCCAGAGAAATCTTCCCGAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360
 QY 361 AATCGTTTGAAGAGTGCAGAGGATCTGCGGCCACCAAAATGATGGGAACAGCTGTGC 420
 Db 361 AATCGTTTGAAGAGTGCAGAGGATCTGCGGCCACCAAAATGATGGGAACAGCTGTGC 420
 QY 421 CCCCCGGGAATCAAGTACCTTCGAGAAGATTAAACAGACATCTGCACCCAAAGGGG 480
 Db 421 CCCCCGGGAATCAAGTACCTTCGAGAAGATTAAACAGACATCTGCACCCAAAGGGG 480
 QY 481 AAACATGCTCGACCCACAGACTCGTGTAGAGAAAGCAGCTGTGTTTATGAGAGATC 540
 Db 481 AAACATGCTCGACCCACAGACTCGTGTAGAGAAAGCAGCTGTGTTTATGAGAGATC 540
 QY 541 AGCCACCTCGAGGAGATGACGACTTACTCCCTCG 576
 Db 541 AGCCACCTCGAGGAGATGACGACTTACTCCCTCG 576

BC005325 1250 bp mRNA linear PRI 03-OCT-2003
 Homo sapiens synovial sarcoma, X breakpoint 4, transcript variant
 1, mRNA (CDNA clone MGC:12411 IMAGE:3961688), complete cds.
 BC005325
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1250)
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.B., Brownstein,M.J., Ueudin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettner,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,I.,
 Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 2338257
 PUBLISHED 12477932
 2 (bases 1 to 1250)
 Strausberg,R.
 Direct Submission
 Submitted (27-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-sngc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>
 Series: IRL Plate: 16 Row: m Column: 20
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 28559011.

FEATURES
 Location/Qualifiers
 1..1250
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:12411 IMAGE:3961688"
 /tissue_type="Bladder, carcinoma"
 /clone_lib="NIH MGC_53"
 /lab_host="DH10B"
 /not5="Vector: pDNR-LIB"
 1..1250
 /gene="SSX4"
 /note="synonym: MGC12411"
 /db_xref="LocustID:6759"
 /db_xref="MIM:300326"
 59..625
 /codon_start=1
 /product="synovial sarcoma, X breakpoint 4, isoform a"
 /protein_id="AAH0325.1"
 /db_xref="GI:13529095"
 /db_xref="LocustID:6759"
 /translation="MKGDDAFRRPRDDAQISEKLKAFDDIAKYFSKKEWKMKSSE
 KIVVMKLNVEVMTKLGKVTLPFPMRSKGAADFGNDGFRNHRVQVPERPQTFG
 SIQRIFPKIPKPAEENGLKEVPEASGPQNDGKQLCPGPNPSTLEKINTKSGPKRG
 KXAWTHLRERKQLVYVEISPEEDDE"
 misc_feature
 125..304
 /note="KXAB; Region: krueppel associated box"
 /db_xref="CDD:smart00349"

ORIGIN
 Query March 99.7%; Score 574.4; DB 9; Length 1250;
 Best Local Similarity 99.8%; Pred. No. 8.8e-164;
 Matches 575; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGAACGGAGACGACGCCCTTTGCAAGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
 Db 59 ATGAACGGAGACGACGCCCTTTGCAAGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 118
 QY 61 TTACGAAAGCCTTCGATGATATGCGCAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
 Db 119 TTACGAAAGCCTTCGATGATATGCGCAATATCTTCTTAAGAAAGAGTGGGAAAGATG 178
 QY 121 AAATCTCTCGAGAAAAATCGTCTATGTATATGAAGCTAAACTATGAGTCTATGACTAA 180
 Db 179 AAATCTCTCGAGAAAAATCGTCTATGTATATGAAGCTAAACTATGAGTCTATGACTAA 238
 QY 181 CTAGGTTTCAAGTTCACCTCCACCTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
 Db 239 CTAGGTTTCAAGTTCACCTCCACCTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 298
 QY 241 GGGAAATGATTTTGTATACGATCGAAACACACAGGAATCAGGTTGAAAGTCTTCAGATGACT 300
 Db 299 GGGAAATGATTTTGTATACGATCGAAACACACAGGAATCAGGTTGAAAGTCTTCAGATGACT 358
 QY 301 TTCCGGCAGCCTCCAGAGAAATCTTCCCGAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360

RESULT 9
 BC005325
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1250)
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.B., Brownstein,M.J., Ueudin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettner,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,I.,
 Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 2338257
 PUBLISHED 12477932
 2 (bases 1 to 1250)
 Strausberg,R.
 Direct Submission
 Submitted (27-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>

TITLE
 JOURNAL
 MEDLINE
 PUBLISHED 12477932
 2 (bases 1 to 1250)
 Strausberg,R.
 Direct Submission
 Submitted (27-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>

REMARK

Db 359 TTGGCAGCCTCCAGAGAAATCTTCCGAGATCATGCCCAAGACGAGAGAGAA 418
QY 361 AATGGTTTGAAGGAGTGCAGAGGATCTGGCCCAACAAAATGATGGGAAACAGCTGTGC 420
Db 419 AATGGTTTGAAGGAGTGCAGAGGATCTGGCCCAACAAAATGATGGGAAACAGCTGTGC 478
QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTACAGACATCTGGACCCAAAAGGGGG 480
Db 479 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTACAGACATCTGGACCCAAAAGGGGG 538
QY 481 AAACATGCTTGGACCCACAGATCTGGTGAAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540
Db 539 AAACATGCTTGGACCCACAGATCTGGTGAAGAAAGCAGCTGGTGGTTTATGAAGAGATC 598
QY 541 AGCAGCCCTGAGAGAGATGACGAGTAACCTCCCTCG 576
Db 599 AGCAGCCCTGAGAGAGATGACGAGTAACCTCCCTCG 634

RESULT 10
A48452
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

A48452
Sequence 17 from Patent WO9602641.
A48452
A48452.1 GI:2302236
unidentified
unclassified.
1 (bases 1 to 766)
Cooper,C.S. and Gusterson,B.A.
MATERIALS AND METHODS RELATING TO THE DIAGNOSIS AND PROPHYLACTIC
AND THERAPEUTIC TREATMENT OF SYNOVIAL SARCOMA
Patent: WO 9602641-A 17 01-FEB-1996;
CANCER RES CAMPAIGN TECH (GB)
Other publication AU 2986595 960216.
Location/Qualifiers
1. .766
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 86.2%; Score 496.6; DB 6; Length 766;
Best Local Similarity 91.5%; Pred. No. 4.6e-140;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATACAGAGAG 60
Db 92 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATACAGAGAG 151
QY 61 TTACGAAAGGCTTCGATGATATTCGCAAAATCTTCTAAGAAAGAGTGGGAAAAGATG 120
Db 152 ATCCAAAAGGCTTCGATGATATTCGCAAAATCTTCTAAGAAAGAGTGGGAAAAGATG 211
QY 121 AAATCTCTCGAGAAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCATGACTAAA 180
Db 212 AAAGCCTCGAGAAAATCTTCTATGTATATGAAGAAAGTATGAGGCTATGACTAAA 271
QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTTTCATCGTAGTAAACGGGCTGCAGACTTCCAC 240
Db 272 CTAGGTTTCAAGGTCAACCTCCACCTTTTCATCGTAGTAAACGGGCTGCAGACTTCCAG 331
QY 241 GGGAAATGATTTGGTAAAGTCCAGAAACCCAGGAATCAGGTTGAACGCTCTCAGATGACT 300
Db 332 GGGAAATGATTTGGTAAAGTCCAGAAACCCAGGAATCAGGTTGAACGCTCTCAGATGACT 391
QY 241 GGGAAATGATTTGGTAAAGTCCAGAAACCCAGGAATCAGGTTGAACGCTCTCAGATGACT 300
Db 332 GGGAAATGATTTGGTAAAGTCCAGAAACCCAGGAATCAGGTTGAACGCTCTCAGATGACT 391
QY 301 TTCCGGAGCTCCAGAGAAATCTTCCGAGATCATGCCAAGATCCAGAGAGCCAGAGAGAGAA 360
Db 392 TTCCGGAGCTCCAGAGAAATCTTCCGAGATCATGCCAAGATCCAGAGAGCCAGAGAGAGAA 451
QY 361 AATGGTTTGAAGGAGTGCAGAGGATCTGGCCCAACAAAATGATGGGAAACAGCTGTGC 420

Db 452 AATGATTCGGAGGAAGTGCAGAGATCTGGCCCAACAAAATGATGGGAAAGAGCTGTGC 511
QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAAGGGGG 480
Db 512 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAAGGGGG 571
QY 481 AAACATGCTTGGACCCACAGATCTGGTGAAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540
Db 572 GAACATGCTTGGACCCACAGATCTGGTGAAGAAAGCAGCTGGTGGTTTATGAAGAGATC 631
QY 541 AGCAGCCCTGAGAGAGATGACGAGTAACCTCCCTCG 575
Db 632 AGCAGCCCTGAGAGAGATGACGAGTAACCTCCCTCG 666

RESULT 11
AX821947
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

AX821947
Sequence 75 from Patent WO03068961.
AX821947
AX821947.1 GI:39725168
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Andrews,P.A., Walsh,J.A. and Gokhale,P.A.
Method to modify differentiation of pluripotential stem cells
Patent: WO 03068961-A 75 21-AUG-2003;
Axordia Limited (GB)
Location/Qualifiers
1. .766
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 86.2%; Score 496.6; DB 6; Length 766;
Best Local Similarity 91.5%; Pred. No. 4.6e-140;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATACAGAGAG 60
Db 92 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATACAGAGAG 151
QY 61 TTACGAAAGGCTTCGATGATATTCGCAAAATCTTCTAAGAAAGAGTGGGAAAAGATG 120
Db 152 ATCCAAAAGGCTTCGATGATATTCGCAAAATCTTCTAAGAAAGAGTGGGAAAAGATG 211
QY 121 AAATCTCTCGAGAAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCATGACTAAA 180
Db 212 AAAGCCTCGAGAAAATCTTCTATGTATATGAAGAAAGTATGAGGCTATGACTAAA 271
QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTTTCATCGTAGTAAACGGGCTGCAGACTTCCAC 240
Db 272 CTAGGTTTCAAGGTCAACCTCCACCTTTTCATCGTAGTAAACGGGCTGCAGACTTCCAG 331
QY 241 GGGAAATGATTTGGTAAAGTCCAGAAACCCAGGAATCAGGTTGAACGCTCTCAGATGACT 300
Db 332 GGGAAATGATTTGGTAAAGTCCAGAAACCCAGGAATCAGGTTGAACGCTCTCAGATGACT 391
QY 301 TTCCGGAGCTCCAGAGAAATCTTCCGAGATCATGCCAAGATCCAGAGAGCCAGAGAGAGAA 360
Db 392 TTCCGGAGCTCCAGAGAAATCTTCCGAGATCATGCCAAGATCCAGAGAGCCAGAGAGAGAA 451
QY 361 AATGGTTTGAAGGAGTGCAGAGGATCTGGCCCAACAAAATGATGGGAAACAGCTGTGC 420
Db 452 AATGATTCGGAGGAAGTGCAGAGATCTGGCCCAACAAAATGATGGGAAAGAGCTGTGC 511
QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAAGGGGG 480
Db 512 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAAGGGGG 571

Db 339 GGGAAATGTTGGATAATGACCCCTAACCGTGGGAATCAGTTTGAACGTCCTCAGATGACT 398
QY 301 TTCGGCAGCTCCAGAGATCTCCGAGATCATGCCAGATCATGCCAGAGCCAGCAGAGAGAA 360
Db 399 TTCGGCAGCTCCAGGAAATCTCCCGAAGATCATGCCAGATCATGCCAGAGCCAGCAGAGAGAA 458
QY 361 AATGGTTTGAAGAAAGTGCAGAGGATCTGCGCCACAAAATGATGGGAAACAGCTGTGC 420
Db 459 AATGATTCGGAGGAAGTGCAGAGATCTGCGCCACAAAATGATGGGAAACAGCTGTGC 518
QY 421 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGG 480
Db 519 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGG 578
QY 481 AAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATC 540
Db 579 GAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATC 638
QY 541 ACGGACCTTGAGAGATCAGAGTAACCTCCCTC 575
Db 639 ACGGACCTTGAGAGATCAGAGTAACCTCCCTC 673

RESULT 14
AR060380
LOCUS AR060380 931 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 2 from patent US 5840568.
ACCESSION AR060380
VERSION AR060380.1 GI:5986830
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 931)
AUTHORS Fireundschuh, M.
TITLE Hodgkin's disease associated molecules and uses thereof
JOURNAL Patent: US 5840568-A 2 24-NOV-1998;
FEATURES
Location/Qualifiers
1..931
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 86.2%; Score 496.6; DB 6; Length 931;
Best Local Similarity 91.5%; Pred. No. 4.7e-140;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGAGACGACCGCTTTCAGAGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
Db 99 ATGAACGAGACGACCGCTTTCAGAGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 158
QY 61 TTACGAAAGGCTTCGATGATATTGCCAAATATCTCTAAGAAAGAGTGGGAAAGATG 120
Db 159 ATCCAAAAGGCTTCGATGATATTGCCAAATATCTCTAAGAAAGAGTGGGAAAGATG 218
QY 121 AAATCTCGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCATGACTAAA 180
Db 219 AAAGCTTCGAGAAATCTTCTATGTATATGAAGAAAGTATGAGGCTATGACTAAA 278
QY 181 CTAGGTTTCAAGTTCACCTCCACCTTCATCGTAGTAAACGGGTGCGAGACTTCCAC 240
Db 279 CTAGGTTTCAAGGCAACCTCCACCTTCATGTATTAACGGCGGAGACTTCCAG 338
QY 241 GGAATGATTTGGTAAACGATCGAAACCCAGGAATCAGGTTGAACGTCCTCAGATGACT 300
Db 339 GGAATGATTTGGTAAATGACCCCTAACCGTGGGAATCAGGTTGAACGTCCTCAGATGACT 398
QY 301 TTGGGAGCTCCAGAGATCTTCCCGAGATCATGCCAGAGATCATGCCAGAGAGAA 360
Db 399 TTGGGAGCTCCAGGAAATCTTCCCGAAGATCATGCCAGAGATCATGCCAGAGAGAA 458
QY 361 AATGGTATTTGGTAAACGATCGAAACCCAGGAATCAGGTTGAACGTCCTCAGATGACT 420
Db 459 AATGATGATTTGGTAAATGACCCCTAACCGTGGGAATCAGGTTGAACGTCCTCAGATGACT 518
QY 421 CCCCCGGGAATCCAGATCTTCCCGAAGATCATGCCAGAGATCATGCCAGAGAGAA 480
Db 519 TTGGGAGCTCCAGGAAATCTTCCCGAAGATCATGCCAGAGATCATGCCAGAGAGAA 578
QY 541 AATGGTATTTGGTAAAGTGCAGAGGATCTGCGCCACAAAATGATGGGAAACAGCTGTGC 600
Db 639 AATGATGATTTGGTAAATGACCCCTAACCGTGGGAATCAGGTTGAACGTCCTCAGATGACT 673

Db 459 AATGATTCGGAGGAAGTGCAGAGATCTGGCCACAAAATGATGGGAAAGAGCTGTGC 518
QY 421 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGG 480
Db 519 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGG 578
QY 481 AAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATC 540
Db 579 GAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATC 638
QY 541 ACGGACCTTGAGAGATGACGAGTAACCTCCCTC 575
Db 639 ACGGACCTTGAGAGATGACGAGTAACCTCCCTC 673

RESULT 15
AR117874
LOCUS AR117874 931 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 2 from patent US 6140464.
ACCESSION AR117874
VERSION AR117874.1 GI:14098780
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 931)
AUTHORS Fireundschuh, M. and Rammensee, H.-G.
TITLE Nonapeptides that bind a HLA-A2.1 molecule
JOURNAL Patent: US 6140464-A 2 31-OCT-2000;
FEATURES
Location/Qualifiers
1..931
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 86.2%; Score 496.6; DB 6; Length 931;
Best Local Similarity 91.5%; Pred. No. 4.7e-140;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGAGACGACCGCTTTCAGAGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
Db 99 ATGAACGAGACGACCGCTTTCAGAGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 158
QY 61 TTACGAAAGGCTTCGATGATATTGCCAAATATCTCTAAGAAAGAGTGGGAAAGATG 120
Db 159 ATCCAAAAGGCTTCGATGATATTGCCAAATATCTCTAAGAAAGAGTGGGAAAGATG 218
QY 121 AAATCTCGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCATGACTAAA 180
Db 219 AAAGCTTCGAGAAATCTTCTATGTATATGAAGAAAGTATGAGGCTATGACTAAA 278
QY 181 CTAGGTTTCAAGTTCACCTCCACCTTCATCGTAGTAAACGGGTGCGAGACTTCCAC 240
Db 279 CTAGGTTTCAAGGCAACCTCCACCTTCATGTATTAACGGCGGAGACTTCCAG 338
QY 241 GGAATGATTTGGTAAACGATCGAAACCCAGGAATCAGGTTGAACGTCCTCAGATGACT 300
Db 339 GGAATGATTTGGTAAATGACCCCTAACCGTGGGAATCAGGTTGAACGTCCTCAGATGACT 398
QY 301 TTGGGAGCTCCAGAGATCTTCCCGAGATCATGCCAGAGATCATGCCAGAGAGAA 360
Db 399 TTGGGAGCTCCAGGAAATCTTCCCGAAGATCATGCCAGAGATCATGCCAGAGAGAA 458
QY 361 AATGGTATTTGGTAAACGATCGAAACCCAGGAATCAGGTTGAACGTCCTCAGATGACT 420
Db 459 AATGATGATTTGGTAAATGACCCCTAACCGTGGGAATCAGGTTGAACGTCCTCAGATGACT 518
QY 421 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGG 480
Db 519 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGG 578
QY 481 AAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATC 540
Db 579 GAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATC 638

Db 579 GAACATGCTGGACCCACAGACTGCGTGAGAGAAAACAGCTGGTGATTATGAAGAGATC 638

QY 541 AGCGACCCCTGAGGAGATGACGAGTAATCTCCCTC 575

Db 639 AGCGACCCCTGAGGAGATGACGAGTAATCTCCCTC 673

Search completed: March 31, 2004, 16:39:15
Job time : 3494 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2004, 13:52:29 ; Search time 437 seconds
(without alignment)
5599.456 Million cell updates/sec

Title: US-09-975-856-1

Perfect score: 576

Sequence: 1 ATGACGGAGACGACGCTT.....ATGACGAGTACTCCCTCG 576

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002s.*
- 7: Geneseqn2003as.*
- 8: Geneseqn2003bs.*
- 9: Geneseqn2003cs.*
- 10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	576	100.0	576	6	ABK84472 Human cDN
2	576	100.0	576	7	ABQ83858 Human SSX
3	576	100.0	1250	9	ADC09598
4	574.4	99.7	576	2	AAV70060 Human SSX
5	496.6	86.2	766	2	AAT11780 Human X-c
6	496.6	86.2	766	6	ABS73282 DNA encod
7	496.6	86.2	766	7	AAD54033 Human col
8	496.6	86.2	766	7	ABQ83844 Human SSX
9	496.6	86.2	766	9	ADC09570
10	496.6	86.2	766	9	ADD25523
11	496.6	86.2	931	2	AAT47748
12	496.6	86.2	931	2	AAV04267 Melanoma
13	496.6	86.2	1309	6	ABL64122 Breast ca
14	496.6	86.2	1309	6	ABL63730 Breast ca
15	487.2	84.6	921	5	AAS80402 DNA encod
16	487.2	84.6	921	5	AAS80399 DNA encod
17	480.6	83.4	576	2	AAV70061 Human SSX
18	480	83.3	766	2	AAT11779 Human X-c
19	221.4	38.4	711	2	AAT11781 Human SVT
20	209.6	36.4	585	2	AAT11782 Human SVT
21	192.6	33.4	830	5	AAS80401 DNA encod
22	191.4	33.2	822	5	AAS80404 DNA encod
23	117.4	20.4	165	6	ABS73281 DNA encod

24	74.2	12.9	1110	5	AAS91978	Aas91978 DNA encod
25	42.2	7.3	2000	7	ADA71938	Ada71938 Rice gene
26	41.4	7.2	1981	9	ADC32171	Adc32171 Human nov
27	40.8	7.1	80	6	ABN33980	Abn33980 Human spl
c 28	40.8	7.1	8059	3	AAAB1747	Aaas1747 N. mening
29	40.8	7.1	110000	3	AAAB1490_07	Continuation (8 of
30	40.8	7.1	349980	3	AAF21608	Aaf21608 Neisseria
c 31	39.8	6.9	571	6	ABK71614	Abk71614 Human dit
c 32	39.8	6.9	3591	8	ADA10966	Ada10966 Human cDN
c 33	39.8	6.9	3591	8	ACH04006	Ach04006 Human cDN
34	39.4	6.8	1097	4	AAF27660	Aaf27660 DNA encod
35	39.4	6.8	1097	4	ACH04007	Ach04007 Human cDN
36	37.8	6.6	2736	4	AAI57976	Aai57976 Human pol
37	37.8	6.6	2872	4	AAH98654	Aah98654 Human EST
38	37.8	6.6	2872	4	AAI59762	Aai59762 Human pol
39	37.8	6.6	5532	9	ADB47398	Adb47398 Human cDN
40	37.4	6.5	730	4	AAC91340	Aac91340 Human pol
41	37	6.4	653	6	ABQ99153	Abq99153 Human ORF
42	36	6.2	2369	6	ABA01156	Aba01156 Human zin
43	36	6.2	2798	9	ADC30277	Adc30277 Human nov
c 44	36	6.2	349980	5	AAH41224	Aah41224 Pyrococcu
45	35.8	6.2	765	4	AAK91889	Aak91889 Human cDN

ALIGNMENTS

RESULT 1
ABK84472
ID ABK84472 standard; cDNA; 576 BP.
XX
AC ABK84472;
XX
DT 14-AUG-2002 (first entry)
XX
DE Human cDNA differentially expressed in granulocytic cells #1043.
XX
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; AKDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX
OS Homo sapiens.
XX
PN WO200228999-A2.
XX
PD 11-APR-2002.
XX
PF 03-OCT-2001; 2001WO-US030821.
XX
PR 03-OCT-2000; 2000US-0237189P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX
XX WPI; 2002-435328/46.
XX
DR
XX
PT Detecting granulocyte activation by detecting differential expression of
genes associated with granulocyte activation, which serves as diagnostic
markers that is useful for monitoring disease states and drug toxicity.
XX
PS Claim 1; SEQ ID NO 1043; 114pp; English.
XX
CC The invention relates to detecting (M1) granulocyte (GC) activation
(GCA), by detecting the level of expression of gene(s) (Gs) identified by
DNA chip analysis as given in the specification, and comparing the
expression level to an expression level in an unactivated GC, where
differential expression of Gs is indicative of GCA. Also included are

modulating (M2) GA by contacting GC with an agent that alters the expression of at least one gene in Gs; (2) screening (M3) for an agent capable of modulating GCA or an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease using the gene expression profile; (3) detecting (M4) an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by detecting the level of expression in a sample of the tissue of gene(s) from Gs, where the level of expression of the gene is indicative of inflammation; (4) treating (M5) an inflammation (especially chronic) or in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease, also bacterial infection, viral infection, parasitic infection, protozoal infection, fungal infection and M5 is useful for treating one of the above conditions. The present sequence represents a gene differentially expressed in granulocytes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 576 BP; 187 A; 127 C; 150 G; 112 T; 0 U; 0 Other;

Query Match 100.0%; Score 576; DB 6; Length 576;
Best Local Similarity 100.0%; Pred. No. 9.3e-172;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGAGAGCGGCTTTCAGAGGAGAGCCAGGATGATGCTCAAAATCAGAGAAG 60
DB 1 ATGACGAGAGCGGCTTTCAGAGGAGAGCCAGGATGATGCTCAAAATCAGAGAAG 60
QY 61 TTACGAAAGGCGCTTCGATGATATTCGCAAAATCTCTTAAGAAAGAGTGGAAGAAGATG 120
DB 61 TTACGAAAGGCGCTTCGATGATATTCGCAAAATCTCTTAAGAAAGAGTGGAAGAAGATG 120
QY 121 AAATCTCGGAGAAATCGTCTATGTGTATGATGAAGCTAACTATGAGTCACTATAA 180
DB 121 AAATCTCGGAGAAATCGTCTATGTGTATGATGAAGCTAACTATGAGTCACTATAA 180
QY 181 CTAGGTTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240
DB 181 CTAGGTTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240
QY 241 GGAATGATTTTGGTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 300
DB 241 GGAATGATTTTGGTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 300
QY 301 TTCGCGAGCTCCAGAGATCTTCGCGAGATCTTCGCGAGATCTTCGCGAGATCTTCGCGAG 360
DB 301 TTCGCGAGCTCCAGAGATCTTCGCGAGATCTTCGCGAGATCTTCGCGAGATCTTCGCGAG 360
QY 361 AATGTTTGAAGAGAGTCCAGAGATCTTCGCGAGATCTTCGCGAGATCTTCGCGAGATCTTCG 420
DB 361 AATGTTTGAAGAGAGTCCAGAGATCTTCGCGAGATCTTCGCGAGATCTTCGCGAGATCTTCG 420
QY 421 CCCCAGGAAATCCAGTACCTTGGAGAGATTTAACAAGATCTTCGAGATCTTCGAGATCTTCGAG 480
DB 421 CCCCAGGAAATCCAGTACCTTGGAGAGATTTAACAAGATCTTCGAGATCTTCGAGATCTTCGAG 480
QY 481 AAACATGCTCGACCCACAGATCTGCGTGAGAGAAAGAGCTGGTGGTGGTGGTGGTGGTGGTGG 540
DB 481 AAACATGCTCGACCCACAGATCTGCGTGAGAGAAAGAGCTGGTGGTGGTGGTGGTGGTGGTGG 540

QY 541 AGGACCGCTCAGGAAGATGACGAGTAACCTCCCTCG 576
DB 541 AGGACCGCTCAGGAAGATGACGAGTAACCTCCCTCG 576

RESULT 2

ABQ83858

ID ABQ83858 standard; cDNA; 576 BP.

AC ABQ83858;

XX 03-FEB-2003 (first entry)

XX Human SSX-4 encoding cDNA SEQ ID NO:599.

XX Human; epitope; vaccine; immunotherapeutic; cytostatic; immunogenicity;

XX T cell; gene; ss.

XX Homo sapiens.

XX WO200281646-A2.

XX 17-OCT-2002.

XX 04-APR-2002; 2002WO-US011101.

XX 06-APR-2001; 2001US-0282211P.

XX 07-NOV-2001; 2001US-0337017P.

XX 07-MAR-2002; 2002US-0363210P.

XX (CTLI-) CTL IMMUNOTHERAPIES CORP.

XX Simard JLL, Diamond DC, Liu L, Xie Z;

XX WPI; 2003-067518/06.

XX P-PSDB; ABP74710.

XX Novel epitopes useful as vaccines, comprises peptides or nucleic acid encoding the peptides, that are useful epitopes of target-associated antigens.

XX Claim 1; Page 185; 352pp; English.

XX The present invention describes an isolated epitope (I) and an epitope cluster. Also described is a vaccine or immunotherapeutic composition (VC) comprising (I). (I) has cytostatic activity. VC is useful for treating an animal, by administering to an animal the vaccine or immunotherapeutic composition. VC is also useful for evaluating immunogenicity of a vaccine or immunotherapeutic composition, by administering VC to an HLA-transgenic animal and evaluating immunogenicity based on a characteristic of the animal, or by in vitro primary stimulation of a T cell and evaluating immunogenicity. (I) is useful for determining specific T cell frequency, by contacting T cells with a MHC-peptide complex, and further comprises ELISPOT analysis, limiting dilution analysis, flow cytometry, in situ hybridisation and/or polymerase chain reaction (PCR). ABQ83858 and ABP74128 to ASP74713 represent sequences used in the exemplification of the present invention

XX Sequence 576 BP; 187 A; 127 C; 150 G; 112 T; 0 U; 0 Other;

Query Match 100.0%; Score 576; DB 7; Length 576;
Best Local Similarity 100.0%; Pred. No. 9.3e-172;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGAGAGAGCGGCTTTCAGAGGAGAGCCAGGATGATGCTCAAAATCAGAGAAG 60
DB 1 ATGACGAGAGAGCGGCTTTCAGAGGAGAGCCAGGATGATGCTCAAAATCAGAGAAG 60
QY 61 TTACGAAAGGCGCTTCGATGATATTCGCAAAATCTCTTAAGAAAGAGTGGAAGAAGATG 120
DB 61 TTACGAAAGGCGCTTCGATGATATTCGCAAAATCTCTTAAGAAAGAGTGGAAGAAGATG 120

QY 121 AAATCTCGGAGAAATCTCTATGTATATGAAGCTAAACTATGAGTTCATGACTAAA 180
DB 121 AAATCTCGGAGAAATCTCTATGTATATGAAGCTAAACTATGAGTTCATGACTAAA 180
QY 181 CTAGGTTTCAAGTTCACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCGAGCTCCAC 240
DB 181 CTAGGTTTCAAGTTCACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCGAGCTCCAC 240
QY 241 GGAATGATTTGGTAAAGTGAACATGAAACACAGAGATCAGTTGAACGTTCCATGACT 300
DB 241 GGAATGATTTGGTAAAGTGAACATGAAACACAGAGATCAGTTGAACGTTCCATGACT 300
QY 301 TTCCGAGCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAACCCAGCAGAGAGAA 360
DB 301 TTCCGAGCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAACCCAGCAGAGAGAA 360
QY 361 AATGTTTGAAGAGTGCAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 420
DB 361 AATGTTTGAAGAGTGCAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 420
QY 421 CCCCCGGAAATCCAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAGGGGG 480
DB 421 CCCCCGGAAATCCAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAGGGGG 480
QY 481 AARATGCTTGGACCCACAGACTCGGTGAGAGAAACAGCTGGTGTATGAAGATC 540
DB 481 AARATGCTTGGACCCACAGACTCGGTGAGAGAAACAGCTGGTGTATGAAGATC 540
QY 541 AGCGACCTGAGAGATGACGAGTAACTCCCTCG 576
DB 541 AGCGACCTGAGAGATGACGAGTAACTCCCTCG 576

RESULT 3
ADC09598
ID ADC09598 standard; cDNA; 1250 BP.

XX AC ADC09598;
XX DT 18-DEC-2003 (first entry)
XX DE SSX-4 cDNA #SEQ ID 599.
XX KW Epitope; immunological; vaccine;
XX KW major histocompatibility complex class I; MHC class I; cancer;
XX KW immunisation; ss.
XX OS Unidentified.
XX PN WO2003008537-A2.
XX PD 30-JAN-2003.
XX PF 29-MAR-2002; 2002WO-US010189.
XX PR 06-APR-2001; 2001US-0282211P.
XX PR 07-NOV-2001; 2001US-0337017P.
XX PR 07-MAR-2002; 2002US-0363210P.
XX PA (CTL1-) CTL IMMUNOTHERAPIES CORP.
XX PI Simard JLL, Diamond DC, Liu L, Xie Z;
XX DR WPI; 2003-248010/24.
XX PT Epitope having high affinity for major histocompatibility complex class I
XX PT useful for treating an animal, evaluating immunogenicity of a vaccine or
XX PT therapeutic composition and for diagnosing a disease.
XX PS Claim 1; SEQ ID NO 599; 239pp; English.
XX CC The invention relates to an isolated epitope polypeptide that has high
XX CC affinity for major histocompatibility complex (MHC) class I, and an

CC epitope cluster comprising the polypeptide. Also disclosed is a vaccine
CC or immunotherapeutic composition containing an epitope of the invention.
CC Compositions of the invention may be used in the treatment of cancer. The
CC method can be combined with a radiation therapy chemotherapy.
CC biochemotherapy or surgery. The composition is also useful for evaluating
CC immunogenicity of a vaccine or immunotherapeutic compound. Multimeric MHC
CC-peptide complexes of the invention are useful for determining specific T
CC cell frequency. This method is useful for evaluating immunological
CC response, by performing the method prior to and subsequent to an
CC immunisation step. Compositions of the invention are useful for
CC diagnosing a disease. The current sequence represents an epitope of the
CC invention with high affinity for MHC class I.

Sequence 1250 BP; 373 A; 302 C; 278 G; 297 T; 0 U; 0 Other;
Query Match 100.0%; Score 576; DB 9; Length 1250;
Best Local Similarity 100.0%; Pred. No. 1.4e-171;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGGATGATCTCAATATCAGAGAAG 60
DB 59 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGGATGATCTCAATATCAGAGAAG 118
QY 61 TTACGAAGGCTTCGATGATATTGCCAATACTTCTTAAGAAAGAGTGGAAAGATG 120
DB 119 TTACGAAGGCTTCGATGATATTGCCAATACTTCTTAAGAAAGAGTGGAAAGATG 178
QY 121 AAATCTCGGAGAAATCGTCTATGTGTATATGAAGCTAAACTATGAGTTCATGACTAAA 180
DB 179 AAATCTCGGAGAAATCGTCTATGTGTATATGAAGCTAAACTATGAGTTCATGACTAAA 238
QY 181 CTAGTTTCAAGGTCACCTCCACCTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
DB 239 CTAGTTTCAAGGTCACCTCCACCTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 298
QY 241 GGAATGATTTGGTAAAGTCCAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 300
DB 299 GGAATGATTTGGTAAAGTCCAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 358
QY 301 TTCCGACCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAACCCAGAGAGAGAGAA 360
DB 359 TTCCGACCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAACCCAGAGAGAGAGAA 418
QY 361 AATGTTTGAAGAGTGCAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 420
DB 419 AATGTTTGAAGAGTGCAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 478
QY 421 CCCCCGGAAATCCAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAGGGGG 480
DB 479 CCCCCGGAAATCCAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAGGGGG 538
QY 481 AARATGCTTGGACCCACAGACTCGGTGAGAGAAACAGCTGGTGTATGAAGATC 540
DB 539 AARATGCTTGGACCCACAGACTCGGTGAGAGAAACAGCTGGTGTATGAAGATC 598
QY 541 AGCGACCTGAGAGATGACGAGTAACTCCCTCG 576
DB 599 AGCGACCTGAGAGATGACGAGTAACTCCCTCG 634

RESULT 4
AAV70060
ID AAV70060 standard; cDNA; 576 BP.
XX AC AAV70060;
XX DT 01-MAR-1999 (first entry)
XX DE Human SSX4 cDNA.
XX KW SSX4; SSX gene; human; tumour associated antigen; cancer; melanoma; ss.
XX OS Homo sapiens.

XX WO9850528-A1.
 XX 12-NOV-1998.
 XX 25-FEB-1998; 98WO-US003661.
 XX 05-MAY-1997; 97US-00851138.
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX Gure AO, Tureci O, Sahin U, Tsang S, Scanlan MJ, Knuth A;
 XX Pfreundschuh M, Old LJ, Chen Y;
 XX WPI; 1998-610379/51.
 XX New SSX gene family members - useful for assaying for cancer cells.
 XX Claim 5; Page 12; 19pp; English.
 XX This is the nucleotide sequence of a newly isolated human SSX4 cDNA
 CC clone. The cDNA was isolated from a human testicular cDNA by PCR
 CC amplification using primers (see AAV70062-63) based on the known SSX2
 CC sequence. A SSX5 clone (see AAV70061) was also obtained. 2 forms of SSX4
 CC were identified. One of these lacked nucleotides 331-456 but was
 CC otherwise identical to the present SSX4 sequence and is described as an
 CC alternatively spliced form. SSX4 shares 89.4% homology to SSX1 on the
 CC nucleotide level and 79.3% at the amino acid level. The inventional
 CC additionally provides expression vectors, transformed or transfected
 CC cells that can be used to produce SSX proteins, and primers (see AAV70062
 CC -73) useful for determining expression of an SSX gene in a sample. The
 CC new SSX genes can be used to assay for cancers such as melanoma
 XX
 XX Sequence 576 BP; 188 A; 127 C; 149 G; 112 T; 0 U; 0 Other;
 SQ
 Query Match 99.7%; Score 574.4; DB 2; Length 576;
 Best Local Similarity 99.8%; Pred. No. 3e-171;
 Matches 575; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGACCGAGACGCGCTTGCAGAGAGACCCAGGATGATCTCAATATCAGAGAAG 60
 DB 1 ATGACCGAGACGCGCTTGCAGAGAGACCCAGGATGATCTCAATATCAGAGAAG 60
 QY 61 TTACGAAAGGCTTCGATGATATGCCAATATCTTCTAAGAAAGAGTGGGAAAGATG 120
 DB 61 TTACGAAAGGCTTCGATGATATGCCAATATCTTCTAAGAAAGAGTGGGAAAGATG 120
 QY 121 AATCCTCGAGAAATCGTCTATGTGTATATGAGCTAACTAAGTCTGAGTCACTAA 180
 DB 121 AATCCTCGAGAAATCGTCTATGTGTATATGAGCTAACTAAGTCTGAGTCACTAA 180
 QY 181 CTAGGTTCAAGGTCACCTCCACCTTTCATGGTATGTAACGGGCTGCAGACTCCAC 240
 DB 181 CTAGGTTCAAGGTCACCTCCACCTTTCATGGTATGTAACGGGCTGCAGACTCCAC 240
 QY 241 GGAATGATTTTGTAGATCGAATACACAGGAATCAGGTTGAACTCTCTCAGATGACT 300
 DB 241 GGAATGATTTTGTAGATCGAATACACAGGAATCAGGTTGAACTCTCTCAGATGACT 300
 QY 301 TTCGCGACCTCCAGAGATCTTCCGAGATCATGCCAAGAGCCAGCAGAGGAGAA 360
 DB 301 TTCGCGACCTCCAGAGATCTTCCGAGATCATGCCAAGAGCCAGCAGAGGAGAA 360
 QY 361 AATGTTTGAAGGAGTCCAGAGGATCTGCGCCACAAAATATGATGGGAAACAGCTGTGC 420
 DB 361 AATGTTTGAAGGAGTCCAGAGGATCTGCGCCACAAAATATGATGGGAAACAGCTGTGC 420
 QY 421 CCCCAGGAATCCAGTACCTTGGAGAGATTACAGACATCTGACCCCAAGGGGG 480
 DB 421 CCCCAGGAATCCAGTACCTTGGAGAGATTACAGACATCTGACCCCAAGGGGG 480
 QY 481 AAACATGCTCGACCCAGAGATCGGTGAGAGAAAGCAGCTGGTGGTTATGAAGATC 540

DB 481 AAACATGCTCGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGGTTATGAAGATC 540
 QY 541 AGCGACCTCAGGAAGATGACGAGTAACCTCCCTCG 576
 DB 541 AGCGACCTCAGGAAGATGACGAGTAACCTCCCTCG 576
 RESULT 5
 AAT11780
 ID AAT11780 standard; cDNA; 766 BP.
 XX AAT11780;
 AC AAT11780;
 DT 16-APR-1996 (first entry)
 XX
 DE Human X-chromosome SSX2 cDNA.
 XX
 KW Human; SSX2 gene; synovial sarcoma; X-chromosome; breakpoint-2;
 KW ornithine-delta-aminotransferase; OATL2 gene; Xp11.2; SSX1 gene;
 KW translocation; chromosome-18; 18q11.2; SYT gene; gene fusion; SYT-SSX2;
 KW fusion protein; primer; PCR; polymerase chain reaction; Smal; LpsI;
 KW probe; antibody; monoclonal antibody; humanised antibody; hybridisation;
 KW antisense; antitumour; recombinant vaccine; vaccinia virus; vector;
 KW cancer; diagnosis; therapy; ss.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT misc_difference 230..232
 FT /tag= a
 FT /codon= seq:ITC, aa:Ser
 FT misc_difference 251..253
 FT /tag= b
 FT /codon= seq:AAG, aa:Asn
 FT misc_difference 257..259
 FT /tag= c
 FT /codon= seq:GAG, aa:Iys
 FT misc_difference 377..379
 FT /tag= d
 FT /codon= seq:CGT, aa:Ileu
 FT misc_recomb 421..422
 FT /tag= e
 FT /note= "Breakpoint for SYT translocation"
 FT misc_feature 422..766
 FT /tag= f
 FT /note= "3'-Region present in SYT-SSX2 fusion"
 FT primer_binding 747..766
 FT /tag= g
 FT /note= "Binds primer AAT11784"
 XX WO9602641-A2.
 PN
 XX
 XX 01-FEB-1996.
 XX
 XX 19-JUL-1995; 95WO-GB001704.
 XX
 XX 19-JUL-1994; 94GB-00014580.
 XX
 XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
 PA
 XX Cooper CS, Gusterson BA;
 PI
 XX WPI; 1996-105904/11.
 DR P-PSDB; AAR90677.
 XX
 XX Diagnosing synovial sarcoma by detection of e.g. SYT and SSX1 sequences,
 PT based on X18 trans-location - by amplification of trans-location
 PT sequences or detection of the encoded protein.
 XX
 XX Claim 6; Fig 7; 55pp; English.
 XX
 XX The sequence represents the SSX2 gene (synovial sarcoma X-chromosome
 CC breakpoint-2). The gene is at the location of a breakpoint at Xp11.2

CC within an ornithine-delta-aminotransferase OATL2 region, associated with
 CC the translocation t(X;18)(p11.2;q11.2) found in human synovial sarcomas,
 CC involving joining of the SYT gene (AAT11778) on chromosome-18 at 18q11.2
 CC to SSX2. The gene fusion is then transcribed to produce an SYT-SSX2
 CC (AAT11781) fusion transcript and translated into a fusion protein. SSX2
 CC may be distinguished from related sequence SSX1 (AAT11779) by digestion
 CC with SmaI and IspI. Primer AAT11784 has been used in polymerase chain
 CC reaction amplification of SSX2 sequences. Detection of the gene or its
 CC product in an abnormal location or as a fusion may be used in diagnosis
 CC of synovial sarcoma, using primers, probes, humanised antibodies,
 CC monoclonal antibodies, etc. Antisense oligonucleotides and antibodies may
 CC also be used therapeutically, and antitumour recombinant vaccines may be
 CC constructed e.g. in a vaccinia virus vector

XX SQ Sequence 766 BP; 229 A; 181 C; 200 G; 156 T; 0 U; 0 Other;

Query Match 86.2%; Score 496.6; DB 2; Length 766;
 Best Local Similarity 91.5%; Pred. No. 1.5e-146;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACACGCTTTGCAAGGAGACCCAGGATGCTCAAAATATCAGAGAAG 60
 DB 92 ATGAACGGAGACACGCTTTGCAAGGAGACCCAGGATGCTCAAAATATCAGAGAAG 151
 QY 61 TTACGAAAGCCCTTCGATGATATGCAAAATCTCTCTAAGAAAGAGTGGAAAGATG 120
 DB 152 ATCAAAAGCCCTTCGATGATATGCAAAATCTCTCTAAGAAAGAGTGGAAAGATG 211
 QY 121 AAATCTCGAGAAATCGTCTATGTATATGAAGCTAAATATGAGTCTATGACTAA 180
 DB 212 AAAGCCCTCGAGAAATCTCTATGTATATGAAGAAAGATGAGGCTATGACTAA 271
 QY 181 CTAGGTTTCAAGTCCACCTCCACCTTCATCGCTAGTAAAGGCTCGAGCTCCAC 240
 DB 272 CTAGGTTTCAAGGCCACCTCCACCTTCATCGTATGTAATAACGGCCGAAAGTCCAG 331
 QY 241 GGAATGATTTTGGTAAACGATGAAACACAGGAATCAGGTTGAACGCTCTCAGATGACT 300
 DB 332 GGAATGATTTTGGTAAATGATGACCTAACCGTGGGAATCAGGTTGAACGCTCTCAGATGACT 391
 QY 301 TTGGGAGCTCCAGAGATCTTCCGAGATCATGCCAGAGATCATGCCAGAGAGAGAA 360
 DB 392 TTGGGAGCTCCAGAGATCTTCCGAGATCATGCCAGAGATCATGCCAGAGAGAGAA 451
 QY 361 AATGGTTTGAAGGAAGTCCAGAGGATCTGGCCCAAAATATGATGGGAAACAGCTGTCC 420
 DB 452 AATGATTCGAGGAGTCCAGAGATCTGGCCCAAAATATGATGGGAAAGCTGTCC 511
 QY 421 CCCCCGGGAAATCCAGTACCTTGGAGAGATTAAACAGACATCTGGACCCCAAAAGGGGG 480
 DB 512 CCCCCGGGAAATCCAGTACCTTGGAGAGATTAAACAGACATCTGGACCCCAAAAGGGGG 571
 QY 481 AAACATGCTGGACCCACAGCTGCGTGAGAGAAAGAGCTGGTGTATGAAGAGATC 540
 DB 572 GACATGCTTGGACCCACAGCTGCGTGAGAGAAAGAGCTGGTGTATGAAGAGATC 631
 QY 541 AGCGACCTCGAGGAAGATGACGAGTAACCTCCCTC 575
 DB 632 AGCGACCTCGAGGAAGATGACGAGTAACCTCCCTC 666

RESULT 6
 ABS73282
 ID ABS73282 standard; DNA; 766 BP.
 XX
 AC ABS73282;
 XX
 DT 04-DEC-2002 (first entry)
 XX
 DE DNA encoding human translocation (X; 18)(p11.2; q11.2) protein #2.
 XX
 KW Chromosome aberration; oncogenic fusion protein; cancer;
 KW proliferative disease; cellular protein isoform; heat shock protein 90;

KW HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;
 KW T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;
 KW acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;
 KW acute lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;
 KW papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;
 KW rhabdomyosarcoma; synovial sarcoma; viral infection; gene; ds.
 XX Homo sapiens.
 OS WO200269900-A2.
 XX 12-SEP-2002.
 PF 01-MAR-2002; 2002WO-US0006518.
 XX 01-MAR-2001; 2001US-0272751P.
 PR (CONF-) CONFORMA THERAPEUTICS CORP.
 PA Fritz LC, Burrows PJ;
 PI WPI; 2002-698710/75.
 DR P-PSDB; ABG95081.
 XX
 PT Treating genetically-defined disease associated with chromosomal
 PT aberrations yielding oncogenic fusion proteins, e.g. cell proliferative
 PT diseases, involves administering an inhibitor of heat shock protein 90.
 XX Disclosure; Page 236; 389pp; English.

The invention describes a method of treating genetically-defined disease associated with chromosomal aberrations yielding oncogenic fusion proteins (I), treating cancerous cells containing (I) in a heterogeneous cell population, treating proliferative diseases associated with mutant protein or cellular protein isoforms (II) dependent on heat shock protein (HSP)-90, or selectively treating cells expressing (II) involving administering HSP90-inhibitor. The method is useful for treating genetically-defined disease with chromosomal aberration yielding oncogenic fusion protein, treating cancerous cells containing fusion protein in heterogeneous cell population, treating proliferative disease (e.g. rheumatoid arthritis or cancer) associated with mutant protein or cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g. p53), or selectively treating cells expressing mutant protein or cellular protein isoform in a patient heterozygous for (II). The method is useful for treating a disease e.g. haematopoietic disorder such as T or B cell lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CMML, or a disease characterised by a solid tumour such as papillary thyroid carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and synovial sarcoma. The method is also useful for treating viral infections. This represents the DNA sequence of a chromosome aberration

Sequence 766 BP; 229 A; 181 C; 200 G; 156 T; 0 U; 0 Other;
 Query Match 86.2%; Score 496.6; DB 6; Length 766;
 Best Local Similarity 91.5%; Pred. No. 1.5e-146;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACACGCTTTGCAAGGAGACCCAGGATGCTCAAAATATCAGAGAAG 60
 DB 92 ATGAACGGAGACACGCTTTGCAAGGAGACCCAGGATGCTCAAAATATCAGAGAAG 151
 QY 61 TTACGAAAGCCCTTCGATGATATGCAAAATCTCTCTAAGAAAGAGTGGAAAGATG 120
 DB 152 ATCAAAAGCCCTTCGATGATATGCAAAATCTCTCTAAGAAAGAGTGGAAAGATG 211
 QY 121 AAATCTCGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCATGACTAAA 180
 DB 212 AAAGCCCTCGAGAAATCTCTATGTATATGAAGAAAGATGAGGCTATGACTAAA 271
 QY 181 CTAGGTTTCAAGGTCACCTCCACCTTCATCGCTAGTAAAGGCTCGAGCTCCAC 240
 DB 272 CTAGGTTTCAAGGTCACCTCCACCTTCATCGTATGTAATAACGGCCGAAAGTCCAG 331

QY 241 GGGATGATTGGTGAACGATCGAAACACAGGAATCAGGTTGAACGCTCTCAGATGACT 300
 Db 332 GGGATGATTGGTGAATGACCTAACCGTGGGAATCAGGTTGAACGCTCTCAGATGACT 391
 QY 301 TTCCGGCAGCTCCAGAGATCTTCCCGAAGATCATGCCCAAGAGCCAGAGGAAGNA 360
 Db 392 TTCCGGCAGGCTCCAGGAATCTCCCGAAGATCATGCCCAAGAGCCAGAGGAAGNA 451
 QY 361 AATGTTTGAAGGAAGTCCAGAGGATCTGGCCCAACAAATGATGGGAAACAGCTGTGC 420
 Db 452 AATGATTGGAGGAAGTCCAGAGGATCTGGCCCAACAAATGATGGGAAACAGCTGTGC 511
 QY 421 CCCCCGGGAATCCAGTACCTTGGAGAGATTACAGACATCTGGACCCCAAGAGGGG 480
 Db 512 CCCCCGGGAATCCAGTACCTTGGAGAGATTACAGACATCTGGACCCCAAGAGGGG 571
 QY 481 AARATGCTCTGGACCCACAGACTGCTGAGAGAAAGCAGCTGGTGTATGAGAGATC 540
 Db 572 GAACATGCTCTGGACCCACAGACTGCTGAGAGAAACAGCTGGTGTATGAGAGATC 631
 QY 541 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTC 575
 Db 632 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTC 666

RESULT 7

AD54033
 ID AAD54033 standard; DNA; 766 BP.

XX AC AAD54033;
 XX

DT 17-JUN-2003 (first entry)
 XX

DE Human colon cancer-associated polypeptide gene, SSX-2.
 XX

KW Human; colon cancer-associated polypeptide; immune response; therapy;
 KW colon cancer; gene; ds.
 XX

OS Homo sapiens.
 XX

PN WO200290986-A1.
 XX

PD 14-NOV-2002.
 XX

XX 02-MAY-2002; 2002WO-US013994.
 XX

XX 04-MAY-2001; 2001US-00849602.
 XX

XX (LUDW-) LUDWIG INST CANCER RES.
 XX (CORR) CORNELL RES FOUND INC.
 PA

XX Chen Y, Old LJ, Scanlan MJ, Stockert E;
 XX WPI; 2003-112003/10.
 XX

XX Diagnosing colon cancer in a subject comprises identifying colon cancer-
 XX associated polypeptides as antigens that elicit immune response in colon
 XX cancer.
 PT

XX Claim 1; Page 90-91; 122pp; English.
 XX

XX The invention relates to a method for diagnosing colon cancer in a
 XX subject which comprises identifying colon cancer-associated polypeptides
 XX as antigens that elicit immune response in colon cancer. The method is
 XX useful for diagnosing, determining onset, progression, or regression of
 XX colon cancer in a subject, or for selecting a course of treatment of a
 XX subject having or suspected of having colon cancer. The colon cancer-
 XX associated polypeptides are useful as markers for diagnosing colon
 XX cancer, and for following the course of treatment of colon cancer. The
 XX present sequence is human colon cancer-associated polypeptide gene
 XX

XX Sequence 766 BP; 229 A; 181 C; 200 G; 156 T; 0 U; 0 Other;
 SQ

Query Match 86.2%; Score 496.6; DB 7; Length 766;
 Best Local Similarity 91.5%; Pred. No. 1.5e-146;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 QY 1 ATGAACGGACGACGCGCTTTGCAAGGAGACCCAGAGATGATGCTCAAAATATCAGAGAAG 60
 Db 92 ATGAACGGACGACGCGCTTTGCAAGGAGACCCAGAGATGATGCTCAAAATATCAGAGAAG 151
 QY 61 TTACGAAGAGCCCTTCGATGATATTGCAAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
 Db 152 ATCAAAAGAGCCCTTCGATGATATTGCAAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 211
 QY 121 AATTCCTCGGAGAAATCGTCTATGTATATCAAGCTAACTATGAGGTCTAGCTATAA 180
 Db 212 AAGCCCTCGGAGAAATCTTCTATGTATATGAAGAAAGTATGAGGTCTAGCTATAA 271
 QY 181 CTAGGTTCAAGGTACCCCTCCACCTTTTCATCGGTAGTAAACGGGTGCGAGACTTCCAC 240
 Db 272 CTAGGTTCAAGGCGACCCCTCCACCTTTTCATCGGTAGTAAACGGGTGCGAGACTTCCAC 331
 QY 241 GGGATGATTGGTGAACGATCGAAGACCCAGAGATGATGCTCAAAATATCAGAGAAG 300
 Db 332 GGGATGATTGGTGAATGATGACCTTAACCGTGGGAATCAGGTTGAACGCTCTCAGATGACT 391
 QY 301 TTCCGGCAGCTCCAGAGATCTTCCCGAAGATCATGCCCAAGAGCCAGAGGAAGNA 360
 Db 392 TTCCGGCAGGCTCCAGGAATCTCCCGAAGATCATGCCCAAGAGCCAGAGGAAGNA 451
 QY 361 AATGTTTGAAGGAAGTCCAGAGGATCTGGCCCAACAAATGATGGGAAACAGCTGTGC 420
 Db 452 AATGATTGGAGGAAGTCCAGAGGATCTGGCCCAACAAATGATGGGAAACAGCTGTGC 511
 QY 421 CCCCCGGGAATCCAGTACCTTGGAGAGATTACAGACATCTGGACCCCAAGAGGGG 480
 Db 512 CCCCCGGGAATCCAGTACCTTGGAGAGATTACAGACATCTGGACCCCAAGAGGGG 571
 QY 481 AAACATGCTGGACCCACAGACTGCTGAGAGAAAGCAGCTGGTGTATGAGAGATC 540
 Db 572 GAACATGCTGGACCCACAGACTGCTGAGAGAAACAGCTGGTGTATGAGAGATC 631
 QY 541 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTC 575
 Db 632 AGCGACCTGAGGAAGATGACGAGTAACCTCCCTC 666

RESULT 8

ABQ83844
 ID ABQ83844 standard; cDNA; 766 BP.

XX AC ABQ83844;
 XX

DT 03-FEB-2003 (first entry)
 XX

DE Human SSX-2 encoding cDNA SEQ ID NO:6.
 XX

KW Human; epitope; vaccine; immunotherapeutic; cytostatic; immunogenicity;
 KW T cell; gene; ss.
 XX

OS Homo sapiens.
 XX

PN WO200281646-A2.
 XX

PD 17-OCT-2002.
 XX

XX 04-APR-2002; 2002WO-US011101.
 XX

XX 06-APR-2001; 2001US-0282211P.
 XX 07-NOV-2001; 2001US-0337017P.
 PR

XX 07-MAR-2002; 2002US-0363210P.
 XX

PA (CTL1-) CTL IMMUNOTHERAPIES CORP.
 XX

PI Simard JLL, Diamond DC, Liu L, Xie Z;

```
XX
DR WPI; 2003-067518/06.
XX P-PSDB; ABP74130.
XX
PT Novel epitopes useful as vaccines, comprises peptides or nucleic acid
PT encoding the peptides, that are useful epitopes of target-associated
PT antigens.
XX
PS Claim 1; Page 148; 352pp; English.
XX
CC The present invention describes an isolated epitope (I) and an epitope
CC cluster. Also described is a vaccine or immunotherapeutic composition
CC (VC) comprising (I). (I) has cytostatic activity. VC is useful for
CC treating an animal, by administering to an animal the vaccine or
CC immunotherapeutic composition. VC is also useful for evaluating
CC immunogenicity of a vaccine or immunotherapeutic composition, by
CC administering VC to an HLA-transgenic animal and evaluating
CC immunogenicity based on a characteristic of the animal, or by in vitro
CC primary stimulation of a T cell and evaluating immunogenicity. (I) is
CC useful for determining specific T cell frequency, by contacting T cells
CC with a MHC-peptide complex, and further comprises ELISPOT analysis,
CC limiting dilution analysis, flow cytometry, in situ hybridisation and/or
CC polymerase chain reaction (PCR). ABQ83843 to ABQ83858 and ABP74128 to
CC ABP741713 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 766 BP; 229 A; 181 C; 200 G; 156 T; 0 U; 0 Other;
Query Match 86.2%; Score 496.6; DB 7; Length 766;
Best Local Similarity 91.5%; Pred. No. 1.5e-146;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
DB 92 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 151
QY 61 TTACGAAAGGCTTCGATGATATGCCAATATCTCTAAGAAAGAGTGGGAAAGATG 120
DB 152 ATCCAAAGGCTTCGATGATATGCCAATATCTCTAAGAAAGAGTGGGAAAGATG 211
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGGTCTATGACTAAA 180
DB 212 AAAGCCTCGGAGAAATCTTCTATGTATATGAAGAAAGATGATGAGGTCTATGACTAAA 271
QY 181 CTAGTTCGAGGTACGCTCCACCTTTCATGCTAGTAAACGGGTGCGAGACTTCCAC 240
Query Match 86.2%; Score 496.6; DB 9; Length 766;
Best Local Similarity 91.5%; Pred. No. 1.5e-146;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
DB 92 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 151
QY 61 TTACGAAAGGCTTCGATGATATGCCAATATCTCTAAGAAAGAGTGGGAAAGATG 120
DB 152 ATCCAAAGGCTTCGATGATATGCCAATATCTCTAAGAAAGAGTGGGAAAGATG 211
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGGTCTATGACTAAA 180
DB 212 AAAGCCTCGGAGAAATCTTCTATGTATATGAAGAAAGATGATGAGGTCTATGACTAAA 271
QY 181 CTAGTTCGAGGTACGCTCCACCTTTCATGCTAGTAAACGGGTGCGAGACTTCCAC 240
```

```
RESULT 9
ADC09570
ID ADC09570 standard; cDNA; 766 BP.
XX
XX ADC09570;
XX
DT 18-DEC-2003 (first entry)
XX
XX SSX-2 cDNA #SEQ ID 6.
XX
XX Epitope; immunological; vaccine;
XX major histocompatibility complex class I; MHC class I; cancer;
XX immunisation; ss.
XX
XX Unidentified.
XX
XX WO2003008537-A2.
XX
XX 30-JAN-2003.
XX
XX 29-MAR-2002; 2002WO-US010189.
XX
XX 06-APR-2001; 2001US-0282211P.
XX
XX 07-NOV-2001; 2001US-0337017P.
XX
XX 07-MAR-2002; 2002US-0383210P.
XX
XX (CTLI-) CTL IMMUNOTHERAPIES CORP.
XX
XX Simard JUL, Diamond DC, Liu L, Xie Z;
XX
XX WPI; 2003-248010/24.
XX
XX Epitope having high affinity for major histocompatibility complex class I
XX useful for treating an animal, evaluating immunogenicity of a vaccine or
XX therapeutic composition and for diagnosing a disease.
XX
XX Claim 1; SEQ ID NO 6; 239pp; English.
XX
XX The invention relates to an isolated epitope polypeptide that has high
XX affinity for major histocompatibility complex (MHC) class I, and an
XX epitope cluster comprising the polypeptide. Also disclosed is a vaccine
XX or immunotherapeutic composition containing an epitope of the invention.
XX Compositions of the invention may be used in the treatment of cancer. The
XX method can be combined with a radiation therapy, chemotherapy,
XX biochemotherapy or surgery. The composition is also useful for evaluating
XX immunogenicity of a vaccine or immunotherapeutic compound. Multimeric MHC
XX -peptide complexes of the invention are useful for determining specific T
XX cell frequency. This method is useful for evaluating immunological
XX response, by performing the method prior to and subsequent to an
XX immunisation step. Compositions of the invention are useful for
XX diagnosing a disease. The current sequence represents an epitope of the
XX invention with high affinity for MHC class I.
XX
XX Sequence 766 BP; 229 A; 181 C; 200 G; 156 T; 0 U; 0 Other;
Query Match 86.2%; Score 496.6; DB 9; Length 766;
Best Local Similarity 91.5%; Pred. No. 1.5e-146;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
DB 92 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 151
QY 61 TTACGAAAGGCTTCGATGATATGCCAATATCTCTAAGAAAGAGTGGGAAAGATG 120
DB 152 ATCCAAAGGCTTCGATGATATGCCAATATCTCTAAGAAAGAGTGGGAAAGATG 211
QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGGTCTATGACTAAA 180
DB 212 AAAGCCTCGGAGAAATCTTCTATGTATATGAAGAAAGATGATGAGGTCTATGACTAAA 271
QY 181 CTAGTTCGAGGTACGCTCCACCTTTCATGCTAGTAAACGGGTGCGAGACTTCCAC 240
```

Db 272 CTAGGTTTCAAGGCCACCTCCACCTTTCATGTGTAATAAACGGGCCGAAAGACTTCCAG 331
 Qy 241 GGGAAATGATTTTGGTAACGATCGAAACCCACAGGAATCAGGTTGAACGCTCTCAGATGACT 300
 Db 332 GGGAAATGATTTGGTAATGATACCTTAACCGTGGGAATCAGGTTGAACGCTCTCAGATGACT 391
 Qy 301 TTGGGAGCCTCCAGAGAAATCTTCCCGAAGATCATGCCAAGAGCCACAGAGAGAGAA 360
 Db 392 TTGGGAGGCTCCAGGAAATCTCCCGAAGATCATGCCAAGAGCCACAGAGAGAGAA 451
 Qy 361 AATGGTTTGAAGAAAGTCCAGAGGATCTTGGCCCAAAATATGATGGGAAACAGCTGTGC 420
 Db 452 AATGATTCGAGGAGTCCAGAGGATCTTGGCCCAAAATATGATGGGAAACAGCTGTGC 511
 Qy 421 CCCCCGGGAATCAAGTACCTTGGAGAGATTTAACAAGACATCTCGACCCAAAAGGGGG 480
 Db 512 CCCCCGGGAATCAAGTACCTTGGAGAGATTTAACAAGACATCTCGACCCAAAAGGGGG 571
 Qy 481 AAACATGCTCGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540
 Db 572 GAACATGCTCGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 631
 Qy 541 AGGACCTCGAGAGATGACGAGTAATCTCCCTC 575
 Db 632 AGGACCTCGAGAGATGACGAGTAATCTCCCTC 666

RESULT 10

ADD25523
 ID ADD25523 standard; DNA; 766 BP.
 AC
 XX ADD25523;
 XX
 DT 15-JAN-2004 (first entry)
 DE Binding domain-immunoglobulin fusion protein-associated DNA #46.
 XX
 KW ds; Binding domain; immunoglobulin; fusion protein; cytostatic;
 KW antiarthritic; immunosuppressive; antidiabetic; antithyroid;
 KW neuroprotective; hinge region; immunoglobulin heavy chain;
 KW CH2 constant region; CH3 constant region; IgG1;
 KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
 KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;
 KW rheumatoid arthritis; myasthenia gravis; Grave's disease;
 KW type I diabetes mellitus; multiple sclerosis; autoimmune disease.
 XX

Unidentified.

OS
 XX US2003118592-A1.

PN
 XX 26-JUN-2003.

PD
 XX 25-JUL-2002; 2002US-00207655.

PF
 XX 17-JAN-2001; 2001US-0367358P.

PR
 XX 17-JAN-2002; 2002US-00053530.

PS
 XX 03-JUN-2002; 2002US-0385691P.

PT
 XX (GENE-) GENE-CRAFT INC.

PI
 XX Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;

PP
 XX WPI; 2003-801317/75.

PT
 XX New binding domain-immunoglobulin fusion protein, useful for treating a

PT
 XX subject having or suspected of having a malignant condition or a B-cell

PT
 XX disorder, e.g. melanoma, Grave's disease or autoimmune disease.

PS
 XX Disclosure; SEQ ID NO 84; 157pp; English.

SC
 XX Unidentified

XX
 XX Sequence 766 BP; 229 A; 181 C; 200 G; 156 T; 0 U; 0 Other;

XX
 XX Pfreundschuh M, Rammensee H;

Query Match 86.2%; Score 496.6; DB 9; Length 766;
 Best Local Similarity 91.5%; Pred. No. 1.5e-146;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 1 ATGAACGGAGACGACGCTTTGCAAGAGACCCAGGATGATCTCAAAATATCAGAGAAG 60
 Db 92 ATGAACGGAGACGACGCTTTGCAAGAGACCCAGGATGATCTCAAAATATCAGAGAAG 151
 Qy 61 TTACGAAGGCTTTGATGATATTGCAAAATCTCTTAAGAAAGAGTGGGAAAAGATG 120
 Db 152 ATCCAAAGGCTTTGATGATATTGCAAAATCTCTTAAGAAAGAGTGGGAAAAGATG 211
 Qy 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGTAACTATGAGTCACTGACTAAA 180
 Db 212 AAAGCTCGGAGAAATCTCTATGTATATGAAGAAAGTATAGGGTATGACTAAA 271
 Qy 181 CTAGGTTTCAAGGTCACCCCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
 Db 272 CTAGGTTTCAAGGTCACCCCTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCAG 331
 Qy 241 GGGAAATGATTTGTAACGATCGAAACCCACAGGAATCAGGTTGAACGCTCTCAGATGACT 300
 Db 332 GGGAAATGATTTGTAACGATCGAAACCCACAGGAATCAGGTTGAACGCTCTCAGATGACT 391
 Qy 301 TTGGGAGCCTCCAGAGAAATCTTCCCGAAGATCATGCCAAGAGCCACAGAGAGAGAA 360
 Db 392 TTGGGAGGCTCCAGAGAAATCTTCCCGAAGATCATGCCAAGAGCCACAGAGAGAGAA 451
 Qy 361 AATGGTTTGAAGAAAGTCCAGAGGATCTTGGCCCAAAATATGATGGGAAACAGCTGTGC 420
 Db 452 AATGATTCGAGGAGTCCAGAGGATCTTGGCCCAAAATATGATGGGAAACAGCTGTGC 511
 Qy 421 CCCCCGGGAATCCAAGTACCTTGGAGAGATTTAACAAGACATCTCGACCCAAAAGGGGG 480
 Db 512 CCCCCGGGAATCCAAGTACCTTGGAGAGATTTAACAAGACATCTCGACCCAAAAGGGGG 571
 Qy 481 AAACATGCTCGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540
 Db 572 GAACATGCTCGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 631
 Qy 541 AGGACCTCGAGAGATGACGAGTAATCTCCCTC 575
 Db 632 AGGACCTCGAGAGATGACGAGTAATCTCCCTC 666

RESULT 11

AAT47748
 ID AAT47748 standard; cDNA; 931 BP.
 AC
 XX AAT47748;
 DT 05-JUN-1997 (first entry)
 DE Melanoma cell antigen HOM-MEL 40 cDNA clone.
 DE Melanoma; tumour antigen; HOM-MEL 40; serological fishing; vaccine; ds.
 KW Homo sapiens.
 OS
 XX WO9640209-A1.
 PN
 XX 19-DEC-1996.
 PD
 XX 07-JUN-1996; 96WO-US009726.
 PF
 XX 07-JUN-1995; 95US-00479328.
 PR
 XX 03-JAN-1996; 96US-00580980.
 PR
 XX 10-MAY-1996; 96US-00644116.
 PS
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA
 XX Pfreundschuh M, Rammensee H;

XX WPI; 1997-051896/05.
 XX Novel method of serological fishing - allows isolation of molecules (esp.
 PT antigens) associated with pathological conditions, e.g. Hodgkin's
 PT disease, cancer or auto-immune disease.
 XX
 XX Claim 31; Page 29; 44pp; English.
 XX

CC A cDNA clone (AA747748) codes for a novel melanoma cell antigen
 CC designated HOM-MEL 40. It was isolated using a novel method in which a
 CC cDNA library prep'd. from malignant melanoma and expressed in E. coli
 CC cells. Lysates of the host cells were screened with sera that had been
 CC treated to remove interfering binding partners. This involved contacting
 CC the sample with lysates of untransfected host cells and with host cells
 CC transformed with the same vector (phage lambda) used to make the cDNA
 CC library. The method, termed serological fishing, can be used to detect
 CC antigens in human tissues, esp. tumour cells, which are useful in the
 CC molecular diagnosis of diseases and/or for immunotherapy and gene therapy
 CC of infectious, autoimmune and malignant diseases (see also AA747747-49).
 CC HOM-MEL 40 is expressed in 50% of melanomas, 20% of prostate cancers, 20%
 CC of gastric cancers, 26% of colorectal cancers, 12% of lung cancers and
 CC 20% of breast cancers, but not in normal tissue. HLA-A2 positive tumour
 CC cells present a nonomer (see also AA09449-52) derived from HOM-MEL,
 CC suggesting that HOM-MEL 40-specific vaccines, useful in inducing
 CC cytotoxic T lymphocytes, are possible
 XX

SQ Sequence 931 BP; 274 A; 207 C; 231 G; 219 T; 0 U; 0 Other;

Query Match 86.2%; Score 496.6; DB 2; Length 931;
 Best Local Similarity 91.5%; Pred. No. 1.6e-146;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 60
 DB 99 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 158
 QY 61 TTACGAAAGGCGCTTCGATGATATTCGCAAAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
 DB 159 ATCCAAAGGCGCTTCGATGATATTCGCAAAATATCTCTTAAGAAAGAGTGGGAAAGATG 218
 QY 121 AAATCCTCGGAGAAATTCGTCTATGTGTATTAAGCTAACTATGAGGTGATGACTTAA 180
 DB 219 AAAGCCTCGGAGAAATTCGTCTATGTGTATTAAGGAAAGATGATGAGGTGATGACTTAA 278
 QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTAGTAAACGGGCTCGAGACTTCCAC 240

RESULT 12

AAV04267
 ID AAV04267 standard; cDNA; 931 BP.
 AC AAV04267;
 XX
 XX 22-JUN-1998 (first entry)
 DT
 XX Melanoma antigen clone HOM-MEL-40.
 DE
 XX Melanoma; marker; antigen; diagnosis; serological fishing; human;
 KW HOM-MEL-40; ss.
 KW Homo sapiens.
 OS
 XX WO9748721-A1.
 FN
 XX 24-DEC-1997.
 PD
 XX 23-JUN-1997; 97WO-US010926.
 PF
 XX 21-JUN-1996; 96US-00668128.
 PR
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA
 XX Pfreundschuh M;
 PI
 XX WPI; 1998-063074/06.
 DR
 XX Nucleic acid and derived protein are markers for Hodgkin's disease - used
 PT in identifying immune reactive markers of disease.
 PT
 XX Example 7; Page 30-31; 47pp; English.

PS Clon HOM-MEL-40 was isolated from a human malignant melanoma cDNA
 CC library using a method, designated serological fishing, designed to
 CC identify immunoreactive markers of disease. In this method, cells
 CC characteristic of a disease are used to prepare a cDNA library for
 CC transformation of eukaryotic or prokaryotic cells, and the cells grown to
 CC express proteins. Patient serum is incubated with the cells used to
 CC prepare the library, but not transfected, to remove any components
 CC reactive with these cells, then the stripped sample is similarly treated
 CC with cells carrying the empty vector. The twice-stripped sample is
 CC incubated with lysate of the library cells, so that specific components
 CC in the sample may bind to the expression protein. Proteins that form
 CC immune complexes are identified as disease markers. In the case of HOM-
 CC MEL-40, the new melanoma associated antigen is strongly expressed in
 CC melanoma, but not healthy tissues. The deduced amino acid sequence
 CC includes 3 tumour-associated peptides (see AA41587-89) that bind to HLA-
 CC A2.1. The serological fishing method was also used to identify a claimed
 CC marker (see AAV04262) of Hodgkin's disease

SQ Sequence 931 BP; 274 A; 207 C; 231 G; 219 T; 0 U; 0 Other;
 Query Match 86.2%; Score 496.6; DB 2; Length 931;
 Best Local Similarity 91.5%; Pred. No. 1.6e-146;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 60
 DB 99 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 158
 QY 61 TTACGAAAGGCGCTTCGATGATATTCGCAAAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
 DB 159 ATCCAAAGGCGCTTCGATGATATTCGCAAAATATCTCTTAAGAAAGAGTGGGAAAGATG 218
 QY 121 AAATCCTCGGAGAAATTCGTCTATGTGTATTAAGCTAACTATGAGGTGATGACTTAA 180
 DB 219 AAAGCCTCGGAGAAATTCGTCTATGTGTATTAAGGAAAGATGATGAGGTGATGACTTAA 278
 QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTAGTAAACGGGCTCGAGACTTCCAC 240

Db 279 CTAGTTTCAAGGCCACCTCCACCTTTTCATGTGTAATAAACGGGCGGAGACTTCCAG 338
 Qy 241 GGAATGATTTGGTAACGATCGAAACACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
 Db 339 GGAATGATTTGGTAATGATACCTTAACCGTGGGAATCAGGTTGAACGTCCTCAGATGACT 398
 Qy 301 TTGGGAGCTCCAGAGATCTTCCCGAGATCATGCCCAAGAGCCGACGAGGAGAA 360
 Db 399 TTGGGAGGCTCAGGAATCTCCCGAAGATCATGCCCAAGAGCCGACGAGGAGAA 458
 Qy 361 AATGTTTGAAGGAATGCGAGGATCTGCGCCCAAAATGATGGGAAACAGCTGTGC 420
 Db 459 AATGTTTGGAGGATGCGAGGATCTGCGCCCAAAATGATGGGAAAGAGCTGTGC 518
 Qy 421 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAAGACATCTCGACCCAAAGGGG 480
 Db 519 CCCCCGGGAATCCAGTACCTTGGAGAGATTAACAAGACATCTCGACCCAAAGGGG 578
 Qy 481 AAACATGCTGAGCCACACAGACTGCGTGAGAGAAAGCAGCTGCTGTTTATGAAGAGATC 540
 Db 579 GAACATGCTGAGCCACACAGACTGCGTGAGAGAAAGCAGCTGCTGTTTATGAAGAGATC 638
 Qy 541 AGCGACCTGAGGAAGATGACGATTAACCTCCCTC 575
 Db 639 AGCGACCTGAGGAAGATGACGATTAACCTCCCTC 673

RESULT 13

ABL64122
 ID ABL64122 standard; DNA; 1309 BP.

XX ABL64122;

XX 15-MAY-2002 (first entry)

XX Breast cancer related gene sequence SEQ ID NO:2459.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 XX cytosolic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 XX gene; ds.

XX Homo sapiens.

XX WO200194629-A2.

XX 13-DEC-2001.

XX 30-MAY-2001; 2001WO-US010838.

XX 05-JUN-2000; 2000US-0209473P.

XX 05-JUN-2000; 2000US-0209531P.

XX 18-SEP-2000; 2000US-0233133P.

XX 18-SEP-2000; 2000US-0233617P.

XX 20-SEP-2000; 2000US-0234009P.

XX 20-SEP-2000; 2000US-0234034P.

XX 20-SEP-2000; 2000US-0234052P.

XX 22-SEP-2000; 2000US-0234509P.

XX 22-SEP-2000; 2000US-0234567P.

XX 25-SEP-2000; 2000US-0234923P.

XX 25-SEP-2000; 2000US-0234924P.

XX 25-SEP-2000; 2000US-0235077P.

XX 25-SEP-2000; 2000US-0235082P.

XX 25-SEP-2000; 2000US-0235134P.

XX 25-SEP-2000; 2000US-0235280P.

XX 26-SEP-2000; 2000US-0235637P.

XX 27-SEP-2000; 2000US-0235638P.

XX 27-SEP-2000; 2000US-0235711P.

XX 27-SEP-2000; 2000US-0235720P.

XX 27-SEP-2000; 2000US-0235840P.

XX 27-SEP-2000; 2000US-0235863P.

XX 28-SEP-2000; 2000US-0236028P.

XX 28-SEP-2000; 2000US-0236032P.

PR 28-SEP-2000; 2000US-0236033P.
 PR 28-SEP-2000; 2000US-0236034P.
 PR 28-SEP-2000; 2000US-0236109P.
 PR 28-SEP-2000; 2000US-0236111P.
 PR 29-SEP-2000; 2000US-0236842P.
 PR 29-SEP-2000; 2000US-0236891P.
 PR 02-OCT-2000; 2000US-0237172P.
 PR 02-OCT-2000; 2000US-0237173P.
 PR 02-OCT-2000; 2000US-0237278P.
 PR 02-OCT-2000; 2000US-0237294P.
 PR 02-OCT-2000; 2000US-0237316P.
 PR 03-OCT-2000; 2000US-0237425P.
 PR 03-OCT-2000; 2000US-0237598P.
 PR 03-OCT-2000; 2000US-0237604P.
 PR 03-OCT-2000; 2000US-0237606P.
 PR 03-OCT-2000; 2000US-0237608P.
 PR 01-NOV-2000; 2000US-0244867P.
 PR 01-NOV-2000; 2000US-0245084P.
 PA (AVAL-) AVALON PHARM.
 XX
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 XX WPI; 2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set.

Claim 1; SEQ ID NO 2459; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 847 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour

Sequence 1309 BP; 385 A; 329 C; 281 G; 309 T; 0 U; 5 Other;

Query Match 86.2%; Score 496.6; DB 6; Length 1309;
 Best Local Similarity 91.5%; Pred. No. 1.9e-146;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 1 ATGACGAGAGCAGCGCTTTCGAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
 Db 99 ATGACGAGAGCAGCGCTTTCGAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 158
 Qy 61 TTACGAAAGGCGCTTCGATGATATTGCCAAATACCTCTCTAAGAAAGAGTGGGAAAGATG 120
 Db 159 ATCCAAAGGCGCTTCGATGATATTGCCAAATACCTCTCTAAGAAAGAGTGGGAAAGATG 218
 Qy 121 AAATCTCGGAGAAATCGTCTATGTATATGAGCTAACTATGAGGTCATGACTATAA 180
 Db 219 AAAGCCCGGAGAAATCTTCTATGTATATGAGGAAAGATGATGAGGTCATGACTATAA 278
 Qy 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCTAGTAGTAAACGGGTCGAGACTTCCAC 240
 Db 279 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCTAGTAGTAAACGGGTCGAGACTTCCAG 338

QY 241 GGGATGATTTTGGTAAAGATCGAATCAAAACACAGGATCAGGTTGACGTCCTCAGATGACT 300
Db 339 GGGATGATTTTGGTAAAGATCGAATCAAAACACAGGATCAGGTTGACGTCCTCAGATGACT 398
QY 301 TTGGCAGCCTCCAGAGATCTTCCGGAAGATCATGCCCCAAGAGCCAGCAGAGAGAGAA 360
Db 399 TTGGCAGCCTCCAGAGATCTTCCGGAAGATCATGCCCCAAGAGCCAGCAGAGAGAGAA 458
QY 361 AATGTTTGAAGGAATGCGCAGAGCATCTGGCCCAAAAATGATGGAAACACGCTGTC 420
Db 459 AATGATTGGAGGAATGCCAGAACATCTGCCCCACAAAATGATGGAAACACGCTGTC 518
QY 421 CCCCCGGGAAATCCAAATACCTTGAGAGAGATTAAAGACATCTGGACCCCAAAAGGGGG 480
Db 519 CCCCCGGGAAATCCAAATACCTTGAGAGAGATTAAAGACATCTGGACCCCAAAAGGGGG 578
QY 481 AAACATGCTGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 540
Db 579 GAACATGCTGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGATC 638
QY 541 AGCGACCTCGAGGAAGATGACGAGTAACCTCCCTC 575
Db 639 AGCGACCTCGAGGAAGATGACGAGTAACCTCCCTC 673

RESULT 14
ID ABL63730
XX ABL63730 standard; DNA; 1309 BP.
AC ABL63730;
XX
DT 15-MAY-2002 (first entry)
XX
DE Breast cancer related gene sequence SEQ ID NO:2067.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 22-SEP-2000; 2000US-0234923P.
PR 23-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 28-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.

PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237808P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 2067; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 1309 BP; 385 A; 329 C; 281 G; 309 T; 0 U; 5 Other;
Query Match 86.2%; Score 496.6; DB 6; Length 1309;
Best Local Similarity 91.5%; Pred. NO. 1.9e-146;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACAGCGCTTTGCAAGGAGACCCAGCGATGCTCAATAATCAGAGAAG 60
Db 99 ATGAACGGAGACAGCGCTTTGCAAGGAGACCCAGCGATGCTCAATAATCAGAGAAG 158
QY 61 TTACGAAAGCGCTTCGATGATTTGCCAAATACTTCTTAAGAAAGAGTGGGAAAAGATG 120
Db 159 ATCCAAAAGCGCTTCGATGATTTGCCAAATACTTCTTAAGAAAGAGTGGGAAAAGATG 218
QY 121 AAATCTCCGAGAAATCGTCTATGTATATGAGCTAAACTATGAGTGCATGACTAA 180
Db 219 AAAGCCTCGGAGAAATCTTCTATGTATATGAGAGAAAGTATGAGGCTATGACTAA 278
QY 181 CTAGGTTTCAAGGTACCCCTCCACCTTCATCGTAGTAAACGGGCTCGAGACTTCCAC 240
Db 279 CTAGGTTTCAAGGCCACCCCTCCACCTTCATGTGTATATGAGAGAAAGTATGAGTAA 338
QY 241 GGGAAATGATTTTGGTAAACGATCGAATCAAAACACAGGATCAGGTTGACGTCCTCAGATGACT 300

Db 339 GGGAAATGTTGGATATGACCTACCTACCGTGGGAATCAGGTTGAACGTCCTCAGATGACT 398
 QY 301 TTCGGAGCCTCCAGAGAAATCTTCCGGAAGATCATGCCAAGAAGCCAGCAGAGGAAGAA 360
 Db 399 TTCGGAGCCTCCAGAGAAATCTTCCGGAAGATCATGCCAAGAAGCCAGCAGAGGAAGAA 458
 QY 361 AATGGTTGAAGAGTCCAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 420
 Db 459 AATGATTCGGAGGAAGTCCAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 518
 QY 421 CCCCCGGGAATCCAAAGTACCTTGGAGAGATTAACAAGACATCTCGACCCCAAGGGGG 480
 Db 519 CCCCCGGGAATCCAAAGTACCTTGGAGAGATTAACAAGACATCTCGACCCCAAGGGGG 578
 QY 481 AATAGCTCTGGAGCCACAGACTGCGTGAGAGAAACAGCTGTGTGTTATGAAGAGATC 540
 Db 579 GAATAGCTCTGGAGCCACAGACTGCGTGAGAGAAACAGCTGTGTGTTATGAAGAGATC 638
 QY 541 AGCGACCTGAGGAAGATGACGAGTAATCTCCCTC 575
 Db 639 AGCGACCTGAGGAAGATGACGAGTAATCTCCCTC 673

RESULT 15

AAS80402
 ID AAS80402 standard; cdna; 921 BP.

AC AAS80402;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #16206.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG16215.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 1; SEQ ID NO 16206; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences
 XX

XX Sequence 921 BP; 267 A; 205 C; 233 G; 216 T; 0 U; 0 Other;

Query Match 84.68; Score 487.2; DB 5; Length 921;
 Best Local Similarity 91.3%; Pred. No. 1.1e-143;
 Matches 528; Conservative 0; Mismatches 48; Indels 2; Gaps 1;

QY 1 ATGAACGGAGCAGCGCTTTGCAAGGAGACCCAGGAGATGATCTCAAAATATCAGAGAAG 60
 Db 54 ATGAACGGAGATGACACCTTTGCAAGGAGACCCAGGCTGGTCTCAATACACAGAGAAG 113
 QY 61 TTACGAAGGCTTCGATGATATTCCTCAATATCTTCTTAAGAAAGAGTGGAAAGATG 120
 Db 114 ATCAAAAGGCTTCGATGATATTCCTCAATATCTTCTTAAGAAAGAGTGGAAAGATG 173
 QY 121 AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCTACACTAAA 180
 Db 174 AAAGTCTCGGAGAAATCGTCTATGTATATGAAGAGAAAGTATGAGGCCATGACTAAA 233
 QY 181 CTAGTTTCAAGGTACCTCCACCTTTCATCGGTAGTAAACGGCTGCAGACTTCCAC 240
 Db 234 CTAGTTTCAAGGCTCCCTCCCATCTTTTCATCGGTAAATAACGGGTTCACAGACTTCCAG 293
 QY 241 GGGAAATGATTTTGGTAAACGATCGAAACACACAGAAATCAGGTTGAACCTCTCAGATGACT 300
 Db 294 GGGATGATTTTGTATATGACCTTAAACCGTGGGAATCAGGTTCAACGCTCTCAGATGACT 353
 QY 301 TTCGGAGCCTCCAGAGAAATCTTCCCGAAGATCATGCCCAAGAACCCAGCAGAGGAGAA 360
 Db 354 TTCGGAGGCTCCAGGGAATCTTCCCGAAGATCATGCCCAAGAACCCAGCAGAGGAGAA 413
 QY 361 AATGTTTGAAGGAAGTCCAGAGGATCTGGCCCAACAAAATGATGGAAACAGCTGTGC 420
 Db 414 AATGTTTGAAGGAAGTCCAGAGGATCTGGCCCAACAAAATGATGGAAACAGCTGTGC 473
 QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAAATTAACAAGACATCTGACCCCAAGAGGGGG 480
 Db 474 CCCCCGGGAAATCCAAAGTACCTTGGAGAAATTAACAAGACATCTGACCCCAAGAGGGGG 533
 QY 481 AAACATGCTGACCCCAAGACTGCGTGAGAGAAAGCAGCTGG - TGGTTTATGAAGAGA 538
 Db 534 GAACATGCTGACCCCAAGACTGCGTGAGAGAAAGCAGCTGGTGGTGGTTTATGAAGAGA 593
 QY 539 TCAGCGACCTTGAGGAAGATGACGAGTAATCTCCCTCG 576
 Db 594 TCAGCGACCTTGAGGAAGATGACGAGTAATCTCCCTTG 631

Search completed: March 31, 2004, 15:40:56

Job time : 441 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2004, 15:19:29 ; Search time 97 Seconds
(without alignments)
3295.378 Million cell updates/sec

Title: US-09-975-856-1

Perfect score: 576

Sequence: 1 ATGAACGGAGACAGCGCTT.....ATGACGAGTAATCCCTCG 576

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/6C_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/6D_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	576	100.0	576	US-09-105-839D-5	Sequence 5, Appli
2	576	100.0	576	US-08-851-138C-1	Sequence 1, Appli
3	576	100.0	576	US-09-796-780-1	Sequence 1, Appli
4	576	100.0	576	US-09-344-040C-5	Sequence 5, Appli
5	576	100.0	576	US-09-833-039A-5	Sequence 5, Appli
6	576	100.0	576	US-09-392-714-17	Sequence 17, Appli
7	496.6	86.2	766	US-09-392-714-16	Sequence 16, Appli
8	496.6	86.2	931	US-08-479-328-2	Sequence 2, Appli
9	496.6	86.2	931	US-08-761-119-2	Sequence 2, Appli
10	496.6	86.2	931	US-08-668-128B-2	Sequence 2, Appli
11	496.6	86.2	931	US-08-905-445-2	Sequence 2, Appli
12	496.6	86.2	931	US-08-959-625-2	Sequence 2, Appli
13	496.6	86.2	931	US-09-008-466-2	Sequence 2, Appli
14	496.6	86.2	931	US-08-580-980A-2	Sequence 2, Appli
15	496.6	86.2	931	US-09-053-453-2	Sequence 2, Appli
16	496.6	86.2	931	US-08-844-116A-2	Sequence 2, Appli
17	496.6	86.2	931	US-09-105-839D-2	Sequence 2, Appli
18	496.6	86.2	931	US-09-437-858-2	Sequence 2, Appli
19	496.6	86.2	931	US-09-344-040C-2	Sequence 2, Appli
20	496.6	86.2	931	US-09-833-039A-2	Sequence 2, Appli
21	482.2	83.7	576	US-09-105-839D-6	Sequence 6, Appli
22	482.2	83.7	576	US-08-851-138C-2	Sequence 2, Appli
23	482.2	83.7	576	US-09-796-780-2	Sequence 2, Appli
24	482.2	83.7	576	US-09-344-040C-6	Sequence 6, Appli
25	482.2	83.7	576	US-09-833-039A-6	Sequence 6, Appli
26	480	83.3	766	US-09-105-839D-1	Sequence 1, Appli
27	480	83.3	766	US-09-344-040C-1	Sequence 1, Appli

c	28	480	83.3	766	4	US-09-833-039A-1	Sequence 1, Appli
	29	50.2	8.7	7218	1	US-08-232-463-14	Sequence 14, Appli
	30	44.2	7.7	448	4	US-09-621-976-9992	Sequence 9992, Ap
	31	39.4	6.8	1097	4	US-09-976-594-115	Sequence 115, App
	32	35.4	6.1	478	4	US-09-621-976-3458	Sequence 3458, Ap
	33	34.6	6.0	2628	4	US-09-294-531B-5	Sequence 5, Appli
	34	33.8	5.9	1185	2	US-08-179-557-14	Sequence 14, Appli
	35	33.4	5.8	5761	4	US-09-620-312D-147	Sequence 147, App
	36	33.4	5.6	6107	4	US-09-976-594-557	Sequence 557, App
	37	32.4	5.6	484	4	US-09-621-976-1092	Sequence 1092, Ap
c	38	32.4	5.6	1001	4	US-09-641-638-257	Sequence 257, App
	39	32.4	5.6	2143	4	US-09-881-578A-3	Sequence 3, Appli
	40	32.2	5.6	358	2	US-08-692-787-5	Sequence 5, Appli
	41	32.2	5.6	358	3	US-09-097-199-5	Sequence 5, Appli
c	42	31.8	5.5	832	4	US-09-621-976-12813	Sequence 2813, Ap
c	43	31.6	5.5	505	4	US-09-621-976-15639	Sequence 15639, A
	44	31.6	5.5	1686	4	US-09-489-039A-3471	Sequence 3471, Ap
c	45	31.6	5.5	1965	4	US-09-489-039A-3333	Sequence 3333, Ap

ALIGNMENTS

RESULT 1

US-09-105-839D-5
; Sequence 5, Application US/09105839D
; Patent No. 6287756
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Sahin, Ugur
; APPLICANT: Gure, Ali
; APPLICANT: Old, Lloyd J
; APPLICANT: Pfeunderschuh, Michael
; TITLE OF INVENTION: Method for Determining Presence of Cancer In A Sample By Deter
; FILE REFERENCE: LUD 5556
; CURRENT APPLICATION NUMBER: US/09/105,839D
; CURRENT FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US 08/851,130
; PRIOR FILING DATE: 1997-05-05
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 5
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-105-839D-5

Query Match	100.0%;	Score 576;	DB 3;	Length 576;
Best Local Similarity	100.0%;	Pred. No. 1.7e-184;		
Mismatches	576;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	ATGAACGGAGACAGCGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG	60	
Db	1	ATGAACGGAGACAGCGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG	60	
Qy	61	TTACGAAAGGCGCTTCGATGATATTCGCAAAATATCTTCTTAAGAAGAGTGGGAAAGATG	120	
Db	61	TTACGAAAGGCGCTTCGATGATATTCGCAAAATATCTTCTTAAGAAGAGTGGGAAAGATG	120	
Qy	121	AAATCCCGGAGAAATCGTCTATGCTATATGAAGTAACTATGAGGTCAATGACTACTAAA	180	
Db	121	AAATCCCGGAGAAATCGTCTATGCTATATGAAGTAACTATGAGGTCAATGACTACTAAA	180	
Qy	181	CTAGGTTTCAAGGTCACCCCTCCACCTTCATGCGTAGTAAACCGGCTGCAGACTTCCAC	240	
Db	181	CTAGGTTTCAAGGTCACCCCTCCACCTTCATGCGTAGTAAACCGGCTGCAGACTTCCAC	240	
Qy	241	GGGAATGATTTTGGTAACGATCGAAACACAGGAATCAGTTGACGTCCTCAGATGACT	300	
Db	241	GGGAATGATTTTGGTAACGATCGAAACACAGGAATCAGTTGACGTCCTCAGATGACT	300	
Qy	301	TTCCGACGCTCCAGAGAATCTTCCCGAAGATCATGCCAAGAACCCAGCAGAGGAAGAA	360	

Db 301 TTCCGACGCTCCAGAGAAATCTTCCGAGAGATCATGCCAAGAGCCAGCAGAGGAAGAA 360
Qy 361 AATGGTTTGAAGAGAGTGCAGAGGATCTGGCCCAAAATATGATGGAAACAGCTGTGC 420
Db 361 AATGGTTTGAAGAGAGTGCAGAGGATCTGGCCCAAAATATGATGGAAACAGCTGTGC 420
Qy 421 CCCCCGGGAATCCAAAGTACCTTGGAGAGATTAACCAAGACATCTGGACCCAAAGGGGG 480
Db 421 CCCCCGGGAATCCAAAGTACCTTGGAGAGATTAACCAAGACATCTGGACCCAAAGGGGG 480
Qy 481 AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540
Db 481 AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540
Qy 541 AGCAGCCTTGAGGAAGATGACGAGTAATCTCCCTCG 576
Db 541 AGCAGCCTTGAGGAAGATGACGAGTAATCTCCCTCG 576

RESULT 2

US-08-851-138C-1
; Sequence 1, Application US/08851138C
; Patent No. 6291658
; GENERAL INFORMATION:
; APPLICANT: Gure, Ali O.; Tureci, Ozlem; Sahin, Ugur; Tsang, Solam;
; APPLICANT: Scanlan, Matthew J.; Knuth Alexander; Pfreundschuh, Michael; Old,
; APPLICANT: Lloyd J.; Chen, Yao-Tseng
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding SSX Family
; TITLE OF INVENTION: Members And Uses Thereof
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,138C
; FILING DATE: 5-May-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6291658man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5480
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3168
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-851-138C-1

Query Match 100.0%; Score 576; DB 3; Length 576;
Best Local Similarity 100.0%; Pred. No. 1.7e-184;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAACGAGACGACGCTTTGCAAGAGAGCCAGGATGATCTCAATATCAGAGAAG 60
Db 1 ATGAACGAGACGACGCTTTGCAAGAGAGCCAGGATGATCTCAATATCAGAGAAG 60
Qy 61 TTACGAAGGCTTCGATGATATTGCCAATATCTCTAAGAAAGAGTGGGAAGATG 120
Db 61 TTACGAAGGCTTCGATGATATTGCCAATATCTCTAAGAAAGAGTGGGAAGATG 120

Qy 121 AAATCCTCGAGAAATCGTCTATGTGTATATGAAGCTAAACATATAGGTCATCACTAAA 180
Db 121 AAATCCTCGAGAAATCGTCTATGTGTATATGAAGCTAAACATATAGGTCATCACTAAA 180
Qy 181 CTAGTTTCAAGGTCAACCTCCACCTTTCATGCTAGTAGTAACCGGCTCGAGACTTCAC 240
Db 181 CTAGTTTCAAGGTCAACCTCCACCTTTCATGCTAGTAGTAACCGGCTCGAGACTTCAC 240
Qy 241 GGGATGATTTTGGTTAAACGATCGAAACCCACAGGAATCAGGTTGAACCTCTCAGATGACT 300
Db 241 GGGATGATTTTGGTTAAACGATCGAAACCCACAGGAATCAGGTTGAACCTCTCAGATGACT 300
Qy 301 TTCGACGCTCCAGAGATCTTCCGAGAGATTCATGCCAAGAGCCAGCAGAGAGAGAA 360
Db 301 TTCGACGCTCCAGAGATCTTCCGAGAGATTCATGCCAAGAGCCAGCAGAGAGAGAA 360
Qy 361 AATGTTTGAAGGAAGTCCAGAGAGCATCTGCCCAACAAATGATGGAAACAGCTGTGC 420
Db 361 AATGTTTGAAGGAAGTCCAGAGAGCATCTGCCCAACAAATGATGGAAACAGCTGTGC 420
Qy 421 CCCCCGGGAATCCAAAGTACCTTGGAGAGATTAACCAAGACATCTGGACCCAAAGGGGG 480
Db 421 CCCCCGGGAATCCAAAGTACCTTGGAGAGATTAACCAAGACATCTGGACCCAAAGGGGG 480
Qy 481 AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540
Db 481 AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGATC 540
Qy 541 AGCAGCCTTGAGGAAGATGACGAGTAATCTCCCTCG 576
Db 541 AGCAGCCTTGAGGAAGATGACGAGTAATCTCCCTCG 576

RESULT 3
US-09-796-780-1
; Sequence 1, Application US/09796780
; Patent No. 6339140
; GENERAL INFORMATION:
; APPLICANT: Gure, Ali O.; Tureci, Ozlem; Sahin, Ugur; Tsang, Solam;
; APPLICANT: Scanlan, Matthew J.; Knuth Alexander; Pfreundschuh, Michael; Old,
; APPLICANT: Lloyd J.; Chen, Yao-Tseng
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding SSX
; TITLE OF INVENTION: Family
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/796,780
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/851,138
; FILING DATE: 5-May-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6339140man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5480
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3168
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 nucleotides

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-796-780-1

Query Match 100.0%; Score 576; DB 4; Length 576;
Best Local Similarity 100.0%; Pred. No. 1.7e-184;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAACGGAGACGACGCCCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
DB 1 ATGAACGGAGACGACGCCCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
QY 61 TTACGAAAGGCCCTTCGATGATATGCAAAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
DB 61 TTACGAAAGGCCCTTCGATGATATGCAAAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
QY 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCAATGCTAA 180
DB 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCAATGCTAA 180
QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTTTCATCGGTAGTAAACGGGCTGCAGACTTCCAC 240
DB 181 CTAGGTTTCAAGGTCAACCTCCACCTTTTCATCGGTAGTAAACGGGCTGCAGACTTCCAC 240
QY 241 GGGATGATTTTGGTAAACGATCGAAACACAGAGATCAGGTTGAACGCTCCTCAGATGACT 300
DB 241 GGGATGATTTTGGTAAACGATCGAAACACAGAGATCAGGTTGAACGCTCCTCAGATGACT 300
QY 301 TTGCGCAGCCTCCAGAGATCTTCCCGAAGATCATGCCAAGAACGACAGAGAGAA 360
DB 301 TTGCGCAGCCTCCAGAGATCTTCCCGAAGATCATGCCAAGAACGACAGAGAGAA 360
QY 361 AATGTTTGAAGAAAGTGCAGAGGATCTGCGCCACAAATATGAGGAAACAGCTGTGC 420
DB 361 AATGTTTGAAGAAAGTGCAGAGGATCTGCGCCACAAATATGAGGAAACAGCTGTGC 420
QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAACAAGACATCTGGACCCAAAAGGGG 480
DB 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAACAAGACATCTGGACCCAAAAGGGG 480
QY 481 AAACATGCTGGACCCACAGATCTGCGTGAGAGAAAGAGCTGGTGTATGAAGATC 540
DB 481 AAACATGCTGGACCCACAGATCTGCGTGAGAGAAAGAGCTGGTGTATGAAGATC 540
QY 541 AGCGACCTCGAGAGATGACGAGTAACCTCCCTCG 576
DB 541 AGCGACCTCGAGAGATGACGAGTAACCTCCCTCG 576

```

RESULT 4

US-09-344-040C-5
Sequence 5, Application US/09344040C
Patent No. 6548064
GENERAL INFORMATION:
APPLICANT: Tureci, Ozlem
APPLICANT: Sahin, Ugur
APPLICANT: Pfreundschuh, Michael
APPLICANT: Ramensschuh, Hans Georg
APPLICANT: Stevanovic, Stefan
TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Determining Expression of an SSX Gene, Peptides Derived From Said SSX Gene a
TITLE OF INVENTION: Gene, and Uses Thereof
FILE REFERENCE: LUD 5556.1
CURRENT FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/344,040C
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/105,839
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: US 08/851,130
NUMBER OF SEQ ID NOS: 132
SEQ ID NO 5
LENGTH: 576
TYPE: DNA
ORGANISM: Homo sapiens

TYPE: DNA
ORGANISM: Homo sapiens
US-09-344-040C-5

Query Match 100.0%; Score 576; DB 4; Length 576;
Best Local Similarity 100.0%; Pred. No. 1.7e-184;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAACGGAGACGACGCCCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
DB 1 ATGAACGGAGACGACGCCCTTTGCAAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
QY 61 TTACGAAAGGCCCTTCGATGATATGCAAAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
DB 61 TTACGAAAGGCCCTTCGATGATATGCAAAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
QY 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCAATGCTAA 180
DB 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCAATGCTAA 180
QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTTTCATCGGTAGTAAACGGGCTGCAGACTTCCAC 240
DB 181 CTAGGTTTCAAGGTCAACCTCCACCTTTTCATCGGTAGTAAACGGGCTGCAGACTTCCAC 240
QY 241 GGGATGATTTTGGTAAACGATCGAAACACAGAGATCAGGTTGAACGCTCCTCAGATGACT 300
DB 241 GGGATGATTTTGGTAAACGATCGAAACACAGAGATCAGGTTGAACGCTCCTCAGATGACT 300
QY 301 TTGCGCAGCCTCCAGAGATCTTCCCGAAGATCATGCCAAGAACGACAGAGAGAA 360
DB 301 TTGCGCAGCCTCCAGAGATCTTCCCGAAGATCATGCCAAGAACGACAGAGAGAA 360
QY 361 AATGTTTGAAGAAAGTGCAGAGGATCTGCGCCACAAATATGAGGAAACAGCTGTGC 420
DB 361 AATGTTTGAAGAAAGTGCAGAGGATCTGCGCCACAAATATGAGGAAACAGCTGTGC 420
QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAACAAGACATCTGGACCCAAAAGGGG 480
DB 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAACAAGACATCTGGACCCAAAAGGGG 480
QY 481 AAACATGCTGGACCCACAGATCTGCGTGAGAGAAAGAGCTGGTGTATGAAGATC 540
DB 481 AAACATGCTGGACCCACAGATCTGCGTGAGAGAAAGAGCTGGTGTATGAAGATC 540
QY 541 AGCGACCTCGAGAGATGACGAGTAACCTCCCTCG 576
DB 541 AGCGACCTCGAGAGATGACGAGTAACCTCCCTCG 576

```

RESULT 5

US-09-833-039A-5
Sequence 5, Application US/09833039A
Patent No. 6673350
GENERAL INFORMATION:
APPLICANT: Tureci, Ozlem
APPLICANT: Sahin, Ugur
APPLICANT: Pfreundschuh, Michael
TITLE OF INVENTION: Tumor Associated Peptide and Uses Thereof
FILE REFERENCE: LUD 5622.1
CURRENT APPLICATION NUMBER: US/09/833,039A
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: US 09/409,455
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/344,040
PRIOR FILING DATE: 1999-06-26
PRIOR APPLICATION NUMBER: US 09/105,839
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 131
SOFTWARE: Patent in version 3.2
SEQ ID NO 5
LENGTH: 576
TYPE: DNA
ORGANISM: Homo sapiens

US-09-833-039A-5

```

Query Match
Best Local Similarity 100.0%; Score 576; DB 4; Length 576;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATGCTCAAAATATCAGAGAAG 60
DB 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATGCTCAAAATATCAGAGAAG 60
QY 61 TTACGAAGGCTTCGATGATATTCGCAAAATATTCCTTAAGAAAGATGGGAAAGATG 120
DB 61 TTACGAAGGCTTCGATGATATTCGCAAAATATTCCTTAAGAAAGATGGGAAAGATG 120
QY 121 AAATCCTCGAGAAAATCGTCTATGTGTATATGAACTAAATATGAGTCTATGACTAAA 180
DB 121 AAATCCTCGAGAAAATCGTCTATGTGTATATGAACTAAATATGAGTCTATGACTAAA 180
QY 181 CTAGGTTTCAGGTCAACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
DB 181 CTAGGTTTCAGGTCAACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
QY 241 GGAATGATTTTGGTAAACGATCGAAGACCAAGGATCAGGTTGAAGTCTCTCAGATGACT 300
DB 241 GGAATGATTTTGGTAAACGATCGAAGACCAAGGATCAGGTTGAAGTCTCTCAGATGACT 300
QY 301 TTCCGAGGCTCCAGAGAATCTTCCCGAAGATCATGCCCAAAATATGAGGAAAGAA 360
DB 301 TTCCGAGGCTCCAGAGAATCTTCCCGAAGATCATGCCCAAAATATGAGGAAAGAA 360
QY 361 AATGGTTTGAAGGAAGTCCAGAGGATCTGCGGCAAGATTAACAGACATCTGGACCCAAAAGGGGG 420
DB 361 AATGGTTTGAAGGAAGTCCAGAGGATCTGCGGCAAGATTAACAGACATCTGGACCCAAAAGGGGG 420
QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGGG 480
DB 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGGG 480
QY 481 AAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
DB 481 AAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
QY 541 AGCGACCTCGAGGAAGATGACGAGTAACCTCCCTCG 576
DB 541 AGCGACCTCGAGGAAGATGACGAGTAACCTCCCTCG 576

```

```

RESULT 6
US-09-392-714-17
; Sequence 17, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE REFERENCE: 10461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; EARLIER FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-392-714-17

```

Query Match 100.0%; Score 576; DB 4; Length 576;

```

Best Local Similarity 100.0%; Pred. No. 1.7e-184;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATGCTCAAAATATCAGAGAAG 60
DB 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATGCTCAAAATATCAGAGAAG 60
QY 61 TTACGAAGGCTTCGATGATATTCGCAAAATATTCCTTAAGAAAGATGGGAAAGATG 120
DB 61 TTACGAAGGCTTCGATGATATTCGCAAAATATTCCTTAAGAAAGATGGGAAAGATG 120
QY 121 AAATCCTCGGAGAAAATCGTCTATGTGTATATGAACTAAATATGAGGTCATGACTAAA 180
DB 121 AAATCCTCGGAGAAAATCGTCTATGTGTATATGAACTAAATATGAGGTCATGACTAAA 180
QY 181 CTAGGTTTCAGGTCAACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
DB 181 CTAGGTTTCAGGTCAACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC 240
QY 241 GGAATGATTTTGGTAAACGATCGAAGACCAAGGATCAGGTTGAAGTCTCTCAGATGACT 300
DB 241 GGAATGATTTTGGTAAACGATCGAAGACCAAGGATCAGGTTGAAGTCTCTCAGATGACT 300
QY 301 TTCCGAGGCTCCAGAGAATCTTCCCGAAGATCATGCCCAAAATATGAGGAAAGAA 360
DB 301 TTCCGAGGCTCCAGAGAATCTTCCCGAAGATCATGCCCAAAATATGAGGAAAGAA 360
QY 361 AATGGTTTGAAGGAAGTCCAGAGGATCTGCGGCAAGATTAACAGACATCTGGACCCAAAAGGGGG 420
DB 361 AATGGTTTGAAGGAAGTCCAGAGGATCTGCGGCAAGATTAACAGACATCTGGACCCAAAAGGGGG 420
QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGGG 480
DB 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAAGGGGG 480
QY 481 AAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
DB 481 AAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
QY 541 AGCGACCTCGAGGAAGATGACGAGTAACCTCCCTCG 576
DB 541 AGCGACCTCGAGGAAGATGACGAGTAACCTCCCTCG 576

```

```

RESULT 7
US-09-392-714-16
; Sequence 16, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE REFERENCE: 10461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; EARLIER FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 766
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-392-714-16

```

Query Match 86.2%; Score 496.6; DB 4; Length 766;
Best Local Similarity 91.5%; Pred. No. 1.3e-157;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCTTTCAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
DB 92 ATGAACGGAGACGACGCTTTCAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 151
QY 61 TTACGAAAGGCTTCGATGATATTCGCAAAATCTCTTAAGAAGAGTGGGAAAGATG 120
DB 152 ATCCAAAAGGCTTCGATGATATTCGCAAAATCTCTTAAGAAGAGTGGGAAAGATG 211
QY 121 AAATCCTCGGAGAAATCGTCTATGTATGAAGCTAAACTATGAGTCTACACTAAA 180
DB 212 AAAGCTCGGAGAAATCTCTATGTATGAAGAGAAATGATGAGGCTATGCTAAA 271
QY 181 CTAGTTTCAAGGTCACCTCCACCTTTTCATGTAGTAAACCGGCTGCAAGTCCAC 240
DB 272 CTAGTTTCAAGGTCACCTCCACCTTTTCATGTATGAAGAGAAATGATGAGGCTATGCTAAA 331
QY 241 GGAATGATTTTGGTAACGATCGAAACACAGGATCAGGTTGAACGCTCCTCAGATGACT 300
DB 332 GGAATGATTTGGTAACGATCGAAACACAGGATCAGGTTGAACGCTCCTCAGATGACT 391
QY 301 TTGCGCAGCTCCAGAGAAATCTTCCGAGAGATCATGCCCAAGAGCCAGAGAGAAAGAA 360
DB 392 TTGCGCAGCTCCAGAGAAATCTTCCGAGAGATCATGCCCAAGAGCCAGAGAGAAAGAA 451
QY 361 AATGCTTTGAAGAGTCCAGAGGATCTGCGCCACAAATGATGGAAACAGTGTGC 420
DB 452 AATGATTCGAGAGAGTCCAGAGGATCTGCGCCACAAATGATGGAAACAGTGTGC 511
QY 421 CCCCCGGGAAATCCAAAGTACCTTTGAGAGAGATTAACAAGACATCTGGACCCCAAAAGGGG 480
DB 512 CCCCCGGGAAATCCAAAGTACCTTTGAGAGAGATTAACAAGACATCTGGACCCCAAAAGGGG 571
QY 481 AAACATGCTGACCCACAGATGCTGAGAGAAAGAGCTGGTGTATGAAGATC 540
DB 572 GAACATGCTGACCCACAGATGCTGAGAGAAAGAGCTGGTGTATGAAGATC 631
QY 541 AGCGACCTCGAGAGAGTACGAGTAACCTCCCTC 575
DB 632 AGCGACCTCGAGAGAGTACGAGTAACCTCCCTC 666

RESULT 8

US-08-479-328-2
; Sequence 2, Application US/08479328
; Patent No. 5698396
; GENERAL INFORMATION:
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Method For Identifying Or Isolating A Molecule
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5698396man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5410
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 931 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
; US-08-479-328-2
Query Match 86.2%; Score 496.6; DB 1; Length 931;
Best Local Similarity 91.5%; Pred. No. 1.5e-157;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACGACGCTTTCAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
DB 99 ATGAACGGAGACGACGCTTTCAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 158
QY 61 TTACGAAAGGCTTCGATGATATTCGCAAAATCTCTTAAGAAGAGTGGGAAAGATG 120
DB 159 ATCCAAAAGGCTTCGATGATATTCGCAAAATCTCTTAAGAAGAGTGGGAAAGATG 218
QY 121 AAATCCTCGGAGAAATCGTCTATGTATGAAGCTAAACTATGAGTCTATGACTAAA 180
DB 219 AAAGCTTCGAGAGAAATCTTCTATGTATGAAGAGAAATGATGAGGCTATGACTAAA 278
QY 181 CTAGTTTCAAGTCCACCTCCACCTTTTCATGCGTAGTAAACGGGCTCAGACTTCCAC 240
DB 279 CTAGTTTCAAGGTCACCTCCACCTTTTCATGCTTAATAACGGGCGAAGACTTCCAG 338
QY 241 GGAATGATTTTGGTAACGATCGAAACACAGGATCAGGTTGAACGCTCCTCAGATGACT 300
DB 339 GGAATGATTTGGTAATGACCTTAACCGTGGGAATCAGGTTGAACGCTCCTCAGATGACT 398
QY 301 TTGCGCAGCTCCAGAGAAATCTTCCGAGAGATCATGCCCAAGAGCCAGAGAGAAAGAA 360
DB 399 TTGCGCAGCTCCAGAGAAATCTTCCGAGAGATCATGCCCAAGAGCCAGAGAGAAAGAA 458
QY 361 AATGCTTTGAAGAGTCCAGAGGATCTGCGCCACAAATGATGGGAAAGAGTGTGC 420
DB 459 AATGATTCGAGAGAGTCCAGAGGATCTGCGCCACAAATGATGGGAAAGAGTGTGC 518
QY 421 CCCCCGGGAAATCCAAAGTACCTTTGAGAGAGATTAACAAGACATCTGGACCCCAAAAGGGG 480
DB 519 CCCCCGGGAAATCCAAAGTACCTTTGAGAGAGATTAACAAGAGATCTGGACCCCAAAAGGGG 578
QY 481 AAACATGCTGACCCACAGATGCTGAGAGAAAGAGCTGGTGTATGAAGATC 540
DB 579 GAACATGCTGACCCACAGATGCTGAGAGAAAGAGCTGGTGTATGAAGATC 638
QY 541 AGCGACCTCGAGAGAGTACGAGTAACCTCCCTC 575
DB 639 AGCGACCTCGAGAGAGTACGAGTAACCTCCCTC 673

RESULT 9

US-08-761-119-2
; Sequence 2, Application US/08761119
; Patent No. 5798264
; GENERAL INFORMATION:
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Method For Identifying Or Isolating A Molecule
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/761,119
 FILING DATE: 6-DECEMBER-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/479,328
 FILING DATE: 07-JUNE-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Hanson, No. 5798264man D.
 REGISTRATION NUMBER: 30,946
 REFERENCE/DOCKET NUMBER: LUD 5410.3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-9200
 TELEFAX: (212) 838-3884
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 931 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double-stranded
 TOPOLOGY: linear
 US-08-761-119-2

Query Match 86.2%; Score 496.6; DB 1; Length 931;
 Best Local Similarity 91.5%; Pred. No. 1.5e-157;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACCCCTTTCGAGGAGACCCAGGATGCTCAATATCAGAGAAG 60
 Db
 QY 99 ATGAACGGAGACGACCCCTTTCGAGGAGACCCAGGATGCTCAATATCAGAGAAG 158
 Db
 QY 61 TTACGAAAGCCCTTCGATGATATTCGAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
 Db
 QY 121 AATCTCTCGAGAAATCGTCTATGTATATGAAGCTAACTATGATGCTATGCTAA 180
 Db
 QY 219 AAGCCTCGAGAAATCTTCTATGTATATGAAGAAAGTATGAGCTATGCTAA 278
 QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTTCATCGTAGTAAACGGGCTGCAGACTCCAC 240
 Db
 QY 279 CTAGGTTTCAAGGACCCCTCCACCTTTTCATGTGTATTAACGGGCGGAGACTCCAG 338
 QY 241 GGAATGATTGGTAAAGATGAAACACAGGATCAGGTTGAACGCTCCTCAGATGACT 300
 Db
 QY 339 GGAATGATTGGTAAATGACCTTAACCGTGGGAATCAGGTTGAACGCTCCTCAGATGACT 398
 QY 301 TTCCGAGGCTCCAGGAAATCTCCGAGATCATGCCGAGATCATGCCGAGAGGAGAA 360
 Db
 QY 399 TTCCGAGGCTCCAGGAAATCTCCGAGATCATGCCGAGAGGAGAA 458
 QY 361 AATGGTTGAAGGAAGTCCAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 420
 Db
 QY 459 AATGATTCGAGGAAGTGCAGAGCATCTGCCCAAAATGATGGGAAAGAGCTGTGC 518
 QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAGGGG 480
 Db
 QY 519 CCCCCGGGAAACCAACTCTCTGAGAGATTCCAGAGATCTGGACCCAAAGGGG 578
 QY 481 AATATGCTCGGACCCAGAGATCTGCTGAGAGAAAGCAGCTGCTGTTATGAAGATC 540
 Db
 QY 579 GAACATGCTCGGACCCAGAGATCTGCTGAGAGAAACAGCTGCTGTTATGAAGATC 638
 QY 541 AGCGACCTCGAGGAAGATCAGGATTAACCTCCCTC 575
 Db 639 AGCGACCTCGAGGAAGATCAGGATTAACCTCCCTC 673

RESULT 10
 US-08-668-128B-2
 Sequence 2, Application US/08668128B
 Patent No. 5840568
 GENERAL INFORMATION:
 APPLICANT: Firends Schuh, Michael
 TITLE OF INVENTION: Hodgkin's Disease Associated Molecules And

TITLE OF INVENTION: Uses Thereof
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Felfe & Lynch
 STREET: 805 Third Avenue
 CITY: New York City
 STATE: New York
 ZIP: 10022
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
 COMPUTER: IBM
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: Wordperfect
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/668,128B
 FILING DATE: 21-JUNE-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/644,116
 FILING DATE: 10-MAY-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/580,980
 FILING DATE: 03-JANUARY-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/479,328
 FILING DATE: 07-JUNE-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Hanson, No. 5840568man D.
 REGISTRATION NUMBER: 30,946
 REFERENCE/DOCKET NUMBER: LUD 5441
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-9200
 TELEFAX: (212) 838-3884
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 931 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double-stranded
 TOPOLOGY: linear
 US-08-668-128B-2

Query Match 86.2%; Score 496.6; DB 2; Length 931;
 Best Local Similarity 91.5%; Pred. No. 1.5e-157;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACCCCTTTCGAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
 Db 99 ATGAACGGAGACGACCCCTTTCGAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 158
 QY 61 TTACGAAAGCCCTTCGATGATATTCGCAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
 Db 159 ATCCAAAGGCCCTTCGATGATATTCGCAATATCTTCTTAAGGAGAGTGGGAAAGATG 218
 QY 121 AATCTCTCGAGAAATCGTCTATGTATATGAAGCTAACTATGATGCTCAGATGACTAA 180
 Db 219 AAGCCTCGAGAAATCTTCTATGTATATGAAGAAAGTATGAGGCTATGACTAA 278
 QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTTCATCGTAGTAAACGGGCTGCAGACTCCAC 240
 Db 279 CTAGGTTTCAAGGACCCCTCCACCTTTTCATGTGTATTAACGGGCGGAGACTCCAG 338
 QY 241 GGAATGATTGGTAAAGATCAGGATTAACCCAGAGAAATCAGGTTGAACGCTCCTCAGATGACT 300
 Db 339 GGAATGATTGGTAAATGACCTTAACCGTGGGAATCAGGTTGAACGCTCCTCAGATGACT 398
 QY 301 TTCCGAGGCTCCAGGAAATCTCCGAGATCATGCCGAGATCATGCCGAGAGGAGAA 360
 Db 399 TTCCGAGGCTCCAGGAAATCTCCGAGATCATGCCGAGAGGAGAA 458
 QY 361 AATGGTTGAAGGAAGTCCAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 420
 Db 459 AATGATTCGAGGAAGTGCAGAGCATCTGCCCAAAATGATGGGAAAGAGCTGTGC 518

LENGTH: 931 base pairs
TYPE: nucleic acid
STRANDEDNESS: double-stranded
TOPOLOGY: linear
US-08-959-625-2

Query Match 86.2%; Score 496.6; DB 3; Length 931;
Best Local Similarity 91.5%; Pred. No. 1.5e-157;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGAGCGCTTTGCAAGGAGACCCAGGATGATCTCAAAATATCAGAGAAG 60
DB 99 ATGAACGGAGACGAGCGCTTTGCAAGGAGACCCAGGATGATCTCAAAATATCAGAGAAG 158
QY 61 TTACGAAAGCGCTTCGATGATATTGCCAAATACCTTCTAAGAAAGAGTGGGAAAGATG 120
DB 159 ATCCAAAGCGCTTCGATGATATTGCCAAATACCTTCTAAGAAAGAGTGGGAAAGATG 218
QY 121 AATCTCTCGAGAAATGCTATGTATATGAAGCTAACTATGAGGTCTAGCTATAA 180
DB 219 AAAGCCTCGAGAAATCTTCTATGTATATGAAGAAAGTATGAGGTCTAGCTATAA 278
QY 181 CTAGGTTTCAAGGTCACCTCCACACCTTTCATGCTAGTAAACGGCTGCAGACTTCCAC 240
DB 279 CTAGGTTTCAAGGTCACCTCCACACCTTTCATGCTAGTAAACGGCTGCAGACTTCCAC 338
QY 241 GGGAAATGTTTGGTAAAGTTCGAACCCAGGAAATCAGGTTGAACGTTCTCAGATGACT 300
DB 339 GGGAAATGTTTGGTAAAGTTCGAACCCAGGAAATCAGGTTGAACGTTCTCAGATGACT 398
QY 301 TTGGGAGCTCCAGAGAAATCTTCCGAGATCATGCCAAGAGCCAGCAGAGGAGAA 360
DB 399 TTGGGAGCTCCAGGAAATCTTCCGAGATCATGCCAAGAGCCAGCAGAGGAGAA 458
QY 361 AATGGTTTGAAGGAAGTTCGAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 420
DB 459 AATGATTCGAGGAAGTTCGAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 518
QY 481 AAACATGCTGGACCCACAGACTGCGTGAAGAAAGCAGCTGGTGTGTTATGAAGATC 540
DB 579 GAACATGCTGGACCCACAGACTGCGTGAAGAAACAGCTGGTGTGTTATGAAGATC 638

RESULT 13

US-09-008-466-2
; Sequence 2, Application US/09008466
; Patent No. 6020134
; GENERAL INFORMATION:
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Method For Identifying Or Isolating A Molecule
; TITLE OF INVENTION: And Molecules Identified Thereby
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/008,466

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/580,980
; FILING DATE: 03-JANUARY-1996
; APPLICATION NUMBER: 08/479,328
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6020134man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5410.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 931 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
; US-09-008-466-2

Query Match 86.2%; Score 496.6; DB 3; Length 931;
Best Local Similarity 91.5%; Pred. No. 1.5e-157;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGAGCGCTTTGCAAGGAGACCCAGGATGATCTCAAAATATCAGAGAAG 60
DB 99 ATGAACGGAGACGAGCGCTTTGCAAGGAGACCCAGGATGATCTCAAAATATCAGAGAAG 158
QY 61 TTACGAAAGCGCTTCGATGATATTGCCAAATACCTTCTAAGAAAGAGTGGGAAAGATG 120
DB 159 ATCCAAAGCGCTTCGATGATATTGCCAAATACCTTCTAAGAAAGAGTGGGAAAGATG 218
QY 121 AATCTCTCGAGAAATGCTATGTATATGAAGCTAACTATGAGGTCTAGCTATAA 180
DB 219 AAAGCCTCGAGAAATCTTCTATGTATATGAAGAAAGTATGAGGTCTAGCTATAA 278
QY 181 CTAGGTTTCAAGGTCACCTCCACACCTTTCATGCTAGTAAACGGCTGCAGACTTCCAC 240
DB 279 CTAGGTTTCAAGGTCACCTCCACACCTTTCATGCTAGTAAACGGCTGCAGACTTCCAC 338
QY 241 GGGAAATGTTTGGTAAAGTTCGAACCCAGGAAATCAGGTTGAACGTTCTCAGATGACT 300
DB 339 GGGAAATGTTTGGTAAAGTTCGAACCCAGGAAATCAGGTTGAACGTTCTCAGATGACT 398
QY 301 TTGGGAGCTCCAGAGAAATCTTCCGAGATCATGCCAAGAGCCAGCAGAGGAGAA 360
DB 399 TTGGGAGCTCCAGGAAATCTTCCGAGATCATGCCAAGAGCCAGCAGAGGAGAA 458
QY 361 AATGGTTTGAAGGAAGTTCGAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 420
DB 459 AATGATTCGAGGAAGTTCGAGAGGATCTGGCCCAAAATGATGGGAAACAGCTGTGC 518
QY 481 AAACATGCTGGACCCACAGACTGCGTGAAGAAAGCAGCTGGTGTGTTATGAAGATC 540
DB 579 GAACATGCTGGACCCACAGACTGCGTGAAGAAACAGCTGGTGTGTTATGAAGATC 638

RESULT 14

US-08-580-980A-2
; Sequence 2, Application US/08580980A
; Patent No. 6025191
; GENERAL INFORMATION:
; APPLICANT: Pfreundschuh, Michael

;; TITLE OF INVENTION: Method For Identifying Or Isolating A Molecule
;; TITLE OF INVENTION: And Molecules Identified Thereby
;; NUMBER OF SEQUENCES: 4
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Felfe & Lynch
;; STREET: 805 Third Avenue
;; CITY: New York City
;; STATE: New York
;; ZIP: 10022
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
;; COMPUTER: IBM
;; OPERATING SYSTEM: PC-DOS
;; SOFTWARE: Wordperfect
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/580,980A
;; FILING DATE: 03-JANUARY-1996
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/479,328
;; FILING DATE: 07-JUNE-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hanson, No. 6025191man D.
;; REGISTRATION NUMBER: 30,946
;; REFERENCE/DOCKET NUMBER: LUD 5410.1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 838-3884
;; TELEFAX: (212) 838-3884
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 931 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double-stranded
;; TOPOLOGY: linear
;; US-08-580-980A-2

Query Match 86.2%; Score 496.6; DB 3; Length 931;
Best Local Similarity 91.5%; Pred. No. 1.5e-157;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACACGCCCTTTGCAAGGAGACCCAGGAGTGTCTCAATATCAGAGAAG 60
DB 99 ATGAACGGAGACACGCCCTTTGCAAGGAGACCCAGGAGTGTCTCAATATCAGAGAAG 158
QY 61 TTACGAAAGCCCTTCGATGATATGCGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
DB 159 ATCCAAAGGCCCTTCGATGATATGCGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 218
QY 121 AAATCTCGGAGAAATCGTCTATGTGTATGAAGCTAAACTATGAGGTCTAGCTATAA 180
DB 219 AAAGCCCTCGGAGAAATCTCTATGTATATGAAGAAAGTATGAGGCTATGACTAAA 278
QY 181 CTAGGTTCAAGGTCACCCCTCCACCTTCATGCGTAGTAAACGGGCTCAGACTTCCAC 240
DB 279 CTAGGTTCAAGGTCACCCCTCCACCTTCATGCGTAGTAAACGGGCTCAGACTTCCAC 338
QY 241 GGGATGATTTTGGTAACGATCGCGTGAGAGAAACAGCTGCTGTTATGAAGATC 540

DB 579 GAACATGCTGTGGACCCACACAGCTGCTGAGAGAAACAGCTGTGATTTATGAAGATC 638
QY 541 AGGACCCCTGAGGAAGATGACGAGTAACCTCCCTC 575
DB 639 AGGACCCCTGAGGAAGATGACGAGTAACCTCCCTC 673

RESULT 15

US-09-053-453-2
; Sequence 2, Application US/09053453
; Patent No. 6103673
; GENERAL INFORMATION:
; APPLICANT: Pfriundschuh, Michael
; TITLE OF INVENTION: Method For Identifying Or Isolating
; TITLE OF INVENTION: A Molecule And Molecules Identified Thereby
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,453
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/580,980
; FILING DATE: 03-JANUARY-1996
; APPLICATION NUMBER: 08/479,328
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6103873man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5410.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 931 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
; US-09-053-453-2

Query Match 86.2%; Score 496.6; DB 3; Length 931;
Best Local Similarity 91.5%; Pred. No. 1.5e-157;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACACGCCCTTTGCAAGGAGACCCAGGAGTGTCTCAATATCAGAGAAG 60
DB 99 ATGAACGGAGACACGCCCTTTGCAAGGAGACCCAGGAGTGTCTCAATATCAGAGAAG 158
QY 61 TTACGAAAGCCCTTCGATGATATGCGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
DB 159 ATCCAAAGGCCCTTCGATGATATGCGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 218
QY 121 AAATCTCGGAGAAATCGTCTATGTGTATGAAGCTAAACTATGAGGTCTAGCTATAA 180
DB 219 AAAGCCCTCGGAGAAATCTCTATGTATATGAAGAAAGTATGAGGCTATGACTAAA 278
QY 181 CTAGGTTCAAGGTCACCCCTCCACCTTCATGCGTAGTAAACGGGCTCAGACTTCCAC 240
DB 279 CTAGGTTCAAGGTCACCCCTCCACCTTCATGCGTAGTAAACGGGCTCAGACTTCCAC 338
QY 241 GGGATGATTTTGGTAACGATCGCGTGAGAGAAACAGCTGCTGTTATGAAGATC 300

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2004, 16:39:19 ; Search time 389 Seconds
(without alignments)
5512.734 Million cell updates/sec

Title: US-09-975-856-1

Perfect score: 576

Sequence: 1 ATGACGAGGACGACGCTT.....ATGACGAGTACTCCCTCG 576

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

```
1: /cgn2_6/prodata/1/pubna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubna/BCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/1/pubna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubna/US09_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/1/pubna/US10A_PUBCOMB.seq:*
14: /cgn2_6/prodata/1/pubna/US10B_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubna/US10C_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubna/US10_NEW_PUB.seq:*
17: /cgn2_6/prodata/1/pubna/US60_NEW_PUB.seq:*
18: /cgn2_6/prodata/1/pubna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	576	100.0	576	10	US-09-975-856-1
2	576	100.0	576	10	US-09-833-039-5
3	576	100.0	576	14	US-10-177-277-5
4	576	100.0	576	15	US-10-177-937-599
5	496.6	86.2	766	10	US-09-849-602-13
6	496.6	86.2	766	14	US-10-207-655-84
7	496.6	86.2	766	15	US-10-177-937-6
8	496.6	86.2	931	10	US-09-833-039-2
9	496.6	86.2	931	14	US-10-177-277-2
10	496.6	86.2	1309	9	US-09-954-531-1000
11	496.6	86.2	1309	9	US-09-954-531-1392
12	496.6	86.1	1249	15	US-10-159-563-402
13	482.2	83.7	576	10	US-09-975-856-2
14	482.2	83.7	576	10	US-09-833-039-6
15	482.2	83.7	576	14	US-10-177-277-6

```
16 480 83.3 766 10 US-09-833-039-1
17 480 83.3 766 14 US-10-177-277-1
18 308.6 53.6 574 15 US-10-027-633-106782
19 308.6 53.6 718 15 US-10-027-633-24594
20 140 24.3 564 14 US-10-029-386-10234
21 137 23.8 189 14 US-10-029-386-23951
22 111.4 19.3 675 15 US-10-027-633-33096
23 111.4 19.3 734 15 US-10-027-633-33097
24 108.6 18.9 734 15 US-10-027-633-148783
25 87.8 15.2 3186778 15 US-10-027-632-174961
26 72 12.5 3225 15 US-10-027-633-114514
27 72 12.5 3225 15 US-10-027-633-114515
28 47 8.2 535 15 US-10-027-633-322431
29 47 8.2 535 15 US-10-027-633-322432
30 41 7.1 583 10 US-09-814-353-18480
31 40.8 7.1 60 10 US-09-908-975-6728
32 39.8 6.9 3591 9 US-09-981-353-84
33 39.6 6.9 439 12 US-10-085-783A-58497
34 39.6 6.9 439 15 US-10-242-535A-58497
35 39.4 6.8 1097 14 US-10-133-013-261
36 39 6.8 629 15 US-10-027-632-134770
37 37.8 6.6 389 12 US-10-085-783A-18965
38 37.8 6.6 389 15 US-10-242-535A-18965
39 37.8 6.6 538 14 US-10-029-386-7375
40 37.8 6.6 1530 10 US-09-814-353-21781
41 37.8 6.6 2024 15 US-10-108-260A-1253
42 37.8 6.6 2736 13 US-10-098-841-181
43 37.8 6.6 3205 14 US-10-269-903-87
44 37.8 6.6 5532 10 US-09-971-394-98
45 37.6 6.5 856 15 US-10-027-632-163719
```

ALIGNMENTS

RESULT 1

```
US-09-975-856-1
; Sequence 1, Application US/09975856
; Publication No. US20030023057A1
; GENERAL INFORMATION:
; APPLICANT: Gure, Ali O.; Tureci, Ozlem; Sahin, Ugur; Teang, Solam;
; Scanlan, Matthew J.; Knuth Alexander; Pfreundschuh, Michael; Old,
; Lloyd J.; Chen, Yao-Tseng
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding SSX
; Family
; Members And Uses Thereof
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/975,856
; FILING DATE: 11-Oct-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/796,780
; FILING DATE: 2001-03-01
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. US20030023057A1man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5480
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3168
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 1:
```

```
Sequence 1, Appli
Sequence 1, Appli
Sequence 106782,
Sequence 24594, A
Sequence 10234, A
Sequence 23951, A
Sequence 33096, A
Sequence 33097, A
Sequence 148783,
Sequence 174961,
Sequence 114514,
Sequence 114515,
Sequence 322431,
Sequence 322432,
Sequence 18480, A
Sequence 6728, Ap
Sequence 84, Appli
Sequence 58497, A
Sequence 58497, A
Sequence 261, App
Sequence 134770,
Sequence 18965, A
Sequence 18965, A
Sequence 7375, Ap
Sequence 21781, A
Sequence 1253, Ap
Sequence 181, App
Sequence 87, Appli
Sequence 98, Appli
Sequence 163719,
```

```

;
; SEQUENCE CHARACTERISTICS: °
; LENGTH: 576 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-975-856-1

Query Match      100.0%; Score 576; DB 10; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.8e-174;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGAACGGAGAGACGACGCGCTTTGCAAGGAGACCCAGGGATGATGCTCAATATACAGAGAAG 60
DB      1  ATGAACGGAGAGACGACGCGCTTTGCAAGGAGACCCAGGGATGATGCTCAATATACAGAGAAG 60

QY      61  TTACGAAGGCGCTTCGATGATATTGCCAAATACTTCTCTAAGAAAGAGTGGGAAAAGATG 120
DB      61  TTACGAAGGCGCTTCGATGATATTGCCAAATACTTCTCTAAGAAAGAGTGGGAAAAGATG 120

QY      121 AAATCCTCGAGAAAATCGTCTATGTGTATGAAGCTAAACTATCAGAGTCATGACTAAA 180
DB      121 AAATCCTCGAGAAAATCGTCTATGTGTATGAAGCTAAACTATCAGAGTCATGACTAAA 180

QY      181 CTAGGTTTCAAGGTCAACCTCCACCTTTTCATGGGTAGTAAACGGGCTCGACAGTTCCAC 240
DB      181 CTAGGTTTCAAGGTCAACCTCCACCTTTTCATGGGTAGTAAACGGGCTCGACAGTTCCAC 240

QY      241 GGGAAATGATTTTGGTAAAGATCGAAAACACAGGAATCAGGTTGAAAGCTCTCAGATGACT 300
DB      241 GGGAAATGATTTTGGTAAAGATCGAAAACACAGGAATCAGGTTGAAAGCTCTCAGATGACT 300

QY      301 TTCCGAGCGCTCCAGAGAATCTTCCGAGAGATCATGCCAAGAGCCAGCAGAGAGAA 360
DB      301 TTCCGAGCGCTCCAGAGAATCTTCCGAGAGATCATGCCAAGAGCCAGCAGAGAGAA 360

QY      361 AATGGTTTGAAGGAAGTCCAGAGAGCATCTGGGCCACAAATGATGGGAAACAGCTGTGC 420
DB      361 AATGGTTTGAAGGAAGTCCAGAGAGCATCTGGGCCACAAATGATGGGAAACAGCTGTGC 420

QY      421 CCCCCGGGAAATCCAAGTACTTTCGGAAGAATTAAACAAGACATCTGGACCCCAAAAGGGG 480
DB      421 CCCCCGGGAAATCCAAGTACTTTCGGAAGAATTAAACAAGACATCTGGACCCCAAAAGGGG 480

QY      481 AAACATGCTCGAGCCACAGACTCGGTGAGAGAAACGCGTGGTGTATTCAACAGATC 540
DB      481 AAACATGCTCGAGCCACAGACTCGGTGAGAGAAACGCGTGGTGTATTCAACAGATC 540

QY      541 AGCGACCTTGAGGAAGATGACGAGTAACTCCCCCTCG 576
DB      541 AGCGACCTTGAGGAAGATGACGAGTAACTCCCCCTCG 576

```

RESULT 2
US-09-833-039-5
; Sequence 5, Application US/09833039
; Publication No. US20030175960A1
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfrendschuh, Michael
; TITLE OF INVENTION: Tumor Associated Peptide and Uses Thereof
; FILE REFERENCE: LUD 5622.1
; CURRENT APPLICATION NUMBER: US/09/833,039
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US 09/409,455
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US 09/344,040
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,839
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US 08/851,130
; PRIOR FILING DATE: 1997-05-05

```

; NUMBER OF SEQ ID NOS: 129
; SEQ ID NO 5
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-039-5

Query Match      100.0%; Score 576; DB 10; Length 576;
Best Local Similarity 100.0%; Pred. No. 2.8e-174;
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy  1  ATGAACGGAGACGAGCGCTTTGCAAGGAGACCCACGGGATGATGCTCAAATATCAGAGAAG 60
Db  1  ATGAACGGAGACGAGCGCTTTGCAAGGAGACCCACGGGATGATGCTCAAATATCAGAGAAG 60

Qy  61  TTACGAAAGGCGCTTCGATGATATTGCCAAATATCTTCTAAGAAAGAGTGGGAAAGATG 120
Db  61  TTACGAAAGGCGCTTCGATGATATTGCCAAATATCTTCTAAGAAAGAGTGGGAAAGATG 120

Qy  121  AAATCTCGGAGAAAATCGTCTATGTGTATATGAAGCTAAACTATGAGTGCATGACTAAA 180
Db  121  AAATCTCGGAGAAAATCGTCTATGTGTATATGAAGCTAAACTATGAGTGCATGACTAAA 180

Qy  181  CTAGGTTTCAAGGTCACCTCCACCTTTTCATCGTAGTAAACGGGCTCGACAGTTCAC 240
Db  181  CTAGGTTTCAAGGTCACCTCCACCTTTTCATCGTAGTAAACGGGCTCGACAGTTCAC 240

Qy  241  GGGAAATGATTTTGTAACGATCGAAACCAACAGGAATCAGGTTGAACGTCTCAGATGACT 300
Db  241  GGGAAATGATTTTGTAACGATCGAAACCAACAGGAATCAGGTTGAACGTCTCAGATGACT 300

Qy  301  TTTCGCGACGCTCCAGAGAAATCTTCCCAAGATCATCCCAAGAGCCACGAGGAAGAA 360
Db  301  TTTCGCGACGCTCCAGAGAAATCTTCCCAAGATCATCCCAAGAGCCACGAGGAAGAA 360

Qy  361  AATGTTTGAAGAGAGTGCACAGAGGCACTCTGGCCCAAAATATGATGGGAAACAGCTGTGC 420
Db  361  AATGTTTGAAGAGAGTGCACAGAGGCACTCTGGCCCAAAATATGATGGGAAACAGCTGTGC 420

Qy  421  CCCC CGGGAATCCAGTACCTTGGAGAGAGATTAAACAGACATCTGGACCCCAAGGGGG 480
Db  421  CCCC CGGGAATCCAGTACCTTGGAGAGAGATTAAACAGACATCTGGACCCCAAGGGGG 480

Qy  481  AAACATGCTGTGACCCACAGACTCGTGGAGAGAAAGCAGCTGGTGGTTTTATGAAGAGATC 540
Db  481  AAACATGCTGTGACCCACAGACTCGTGGAGAGAGAAAGCAGCTGGTGGTTTTATGAAGATC 540

Qy  541  AGCGACCTCGAGGAAGATGACGAGTAACCTCCCTCG 576
Db  541  AGCGACCTCGAGGAAGATGACGAGTAACCTCCCTCG 576

```

```

RESULT 3
US-10-177-277-5
; Sequence 5, Application US/10177277
; Publication No. US20030185844A1
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; APPLICANT: Ramensee, Hans Georg
; APPLICANT: Stevanovic, Stefan
; TITLE OF INVENTION: Method for Determining Presence of Cancer In a Sample By Detecting Expression of an Ssx Gene, Peptides Derived From Said Ssx Gene, and Uses Thereof
; TITLE OF INVENTION: Expression of an Ssx Gene, Peptides Derived From Said Ssx Gene, and Uses Thereof
; FILE REFERENCE: LUD 5556.1
; CURRENT APPLICATION NUMBER: US/10/177,277
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US/09/344,040
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,839
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US 08/851,130

```

; PRIOR FILING DATE: 1997-05-05
 ; NUMBER OF SEQ ID NOS: 132
 ; SEQ ID NO 5
 ; LENGTH: 576
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-177-277-5

Query Match 100.0%; Score 576; DB 14; Length 576;
 Best Local Similarity 100.0%; Pred. No. 2.8e-174;
 Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAACGGAGACGACGCTTTCGAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAG	60
DB	1	ATGAACGGAGACGACGCTTTCGAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAG	60
QY	61	TTACGAAAGGCTTCGATGATATTCGCAAAATATCTCTAAGAAAGAGTGGGAAAGATG	120
DB	61	TTACGAAAGGCTTCGATGATATTCGCAAAATATCTCTAAGAAAGAGTGGGAAAGATG	120
QY	121	AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCACTACTAAA	180
DB	121	AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCACTACTAAA	180
QY	181	CTAGGTTTCAAGTCCACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC	240
DB	181	CTAGGTTTCAAGTCCACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC	240
QY	241	GGGAATGATTTTGGTAAACGATCGAAACACACAGGAATCAGGTTGAACGTCCTCAGATGACT	300
DB	241	GGGAATGATTTTGGTAAACGATCGAAACACACAGGAATCAGGTTGAACGTCCTCAGATGACT	300
QY	301	TTCCGACGCTCCAGAGATCTTCCGAGATCATGCCAGAGATCATGCCAGAGCAGAGAGAGAA	360
DB	301	TTCCGACGCTCCAGAGATCTTCCGAGATCATGCCAGAGATCATGCCAGAGCAGAGAGAGAA	360
QY	361	AATGTTTGAAGGAAGTGCAGAGGATCTGCCACACAAAATGATGGAAACAGCTGTGC	420
DB	361	AATGTTTGAAGGAAGTGCAGAGGATCTGCCACACAAAATGATGGAAACAGCTGTGC	420
QY	421	CCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG	480
DB	421	CCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG	480
QY	481	AAACATGCTGCACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATC	540
DB	481	AAACATGCTGCACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATC	540
QY	541	AGCGACCTTGAGGAAGATGACGATTAACCTCCCTCG	576
DB	541	AGCGACCTTGAGGAAGATGACGATTAACCTCCCTCG	576

RESULT 4
 US-10-117-937-599
 ; Sequence 599, Application US/10117937
 ; Publication No. US20030220239A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CTL IMMUNO THERAPIES CORP.
 ; APPLICANT: SIMARD, John, J.L.
 ; APPLICANT: DIAMOND, David, C.
 ; APPLICANT: LIU, Liping
 ; APPLICANT: XIE, Zhidong
 ; TITLE OF INVENTION: EPITOPE SEQUENCES
 ; FILE REFERENCE: CTLLMM.027A
 ; CURRENT APPLICATION NUMBER: US/10/117,937
 ; CURRENT FILING DATE: 2002-04-04
 ; PRIOR APPLICATION NUMBER: US 60/282,211
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: US 60/337,017
 ; PRIOR FILING DATE: 2001-11-07
 ; PRIOR APPLICATION NUMBER: US 60/363,210
 ; PRIOR FILING DATE: 2002-03-07

; NUMBER OF SEQ ID NOS: 602
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 599
 ; LENGTH: 576
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-117-937-599

Query Match 100.0%; Score 576; DB 15; Length 576;
 Best Local Similarity 100.0%; Pred. No. 2.8e-174;
 Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAACGGAGACGACGCTTTCGAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAG	60
DB	1	ATGAACGGAGACGACGCTTTCGAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAG	60
QY	61	TTACGAAAGGCTTCGATGATATTCGCAAAATATCTCTAAGAAAGAGTGGGAAAGATG	120
DB	61	TTACGAAAGGCTTCGATGATATTCGCAAAATATCTCTAAGAAAGAGTGGGAAAGATG	120
QY	121	AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCACTACTAAA	180
DB	121	AAATCCTCGGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCACTACTAAA	180
QY	181	CTAGGTTTCAAGTCCACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC	240
DB	181	CTAGGTTTCAAGTCCACCTCCACCTTTTCATGCGTAGTAAACGGGCTGCAGACTTCCAC	240
QY	241	GGGAATGATTTTGGTAAACGATCGAAACACACAGGAATCAGGTTGAACGTCCTCAGATGACT	300
DB	241	GGGAATGATTTTGGTAAACGATCGAAACACACAGGAATCAGGTTGAACGTCCTCAGATGACT	300
QY	301	TTCCGACGCTCCAGAGATCTTCCGAGATCATGCCAGAGATCATGCCAGAGCAGAGAGAGAA	360
DB	301	TTCCGACGCTCCAGAGATCTTCCGAGATCATGCCAGAGATCATGCCAGAGCAGAGAGAGAA	360
QY	361	AATGTTTGAAGGAAGTGCAGAGGATCTGCCACACAAAATGATGGAAACAGCTGTGC	420
DB	361	AATGTTTGAAGGAAGTGCAGAGGATCTGCCACACAAAATGATGGAAACAGCTGTGC	420
QY	421	CCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG	480
DB	421	CCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG	480
QY	481	AAACATGCTGCACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATC	540
DB	481	AAACATGCTGCACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATTAAGAGATC	540
QY	541	AGCGACCTTGAGGAAGATGACGATTAACCTCCCTCG	576
DB	541	AGCGACCTTGAGGAAGATGACGATTAACCTCCCTCG	576

RESULT 5
 US-09-849-602-13
 ; Sequence 13, Application US/09849602
 ; Publication No. US20030165834A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Scanlan, Matthew J.
 ; APPLICANT: Old, Lloyd J.
 ; APPLICANT: Stockert, Elisabeth
 ; APPLICANT: Chen, Yao-Tseng
 ; TITLE OF INVENTION: Colon Cancer Antigen Panel
 ; FILE REFERENCE: L0461/7105(JRV)
 ; CURRENT APPLICATION NUMBER: US/09/849,602
 ; CURRENT FILING DATE: 2001-05-04
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 13
 ; LENGTH: 766
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-849-602-13

QY 61 TTACGAAGGCTTCGATGATATTCGCAAAATCTCTCTAAGAAGAGTGGGAAAAGATG 120
DB 152 ATCCAAAAGGCTTCGATGATATTCGCAAAATCTCTCTAAGAAGAGTGGGAAAAGATG 211
QY 121 AAATCCTCGAGAAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCTAGCTAAA 180
DB 212 AAAGCCTCGAGAAAATCTTCTATGTATATGAAGAAAGTATGAGGCTATGACTAAA 271
QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTCCTATGCTAGTAAACCGGCTGCGAGCTTCCAC 240
DB 272 CTAGGTTTCAAGGCCACCTCCACCTTCCTATGCTAGTAAACCGGCTGCGAGCTTCCAG 331
QY 241 GGAATGATTTTGGTAAACGATCGAAACACACAGGAATCAGGTTTGAACCTCTCAGATGACT 300
DB 332 GGAATGATTTGGTAAACGATCGAAACACACAGGAATCAGGTTTGAACCTCTCAGATGACT 391
QY 301 TTGGGAGCCTCCAGAAATCTTCCGAAAGTATGCGCAAGTATGCGCAAGCCAGAGAGAA 360
DB 392 TTGGGAGCCTCCAGAAATCTTCCGAAAGTATGCGCAAGTATGCGCAAGCCAGAGAGAA 451
QY 361 AATGTTTGAAGGAGTGCAGAGGATCTGCGCAAGTATGCGCAAGTATGCGCAAGCCAGAGAG 420
DB 452 AATGATTTGGGAGGAGTGCAGAGGATCTGCGCAAGTATGCGCAAGTATGCGCAAGCCAGAGAG 511
QY 421 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTTAAACAGACATCTGGACCCCAAAAGGGGG 480
DB 512 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTTAAACAGACATCTGGACCCCAAAAGGGGG 571
QY 481 AAACATGCTGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATGAGAGATC 540
DB 572 GAAATGCTGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATGAGAGATC 631
QY 541 AGCGACCTTGAGGAAGATGACGAGTAACCTCCCTC 575
DB 632 AGCGACCTTGAGGAAGATGACGAGTAACCTCCCTC 666

RESULT 8

US-09-833-039-2
; Sequence 2, Application US/09833039
; Publication No. US20030175960A1
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Tumor Associated Peptide and Uses Thereof
; FILE REFERENCE: LUD 5622.1
; CURRENT APPLICATION NUMBER: US/09/833,039
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: US/09/409,455
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US/09/344,040
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US/09/105,839
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US/08/851,130
; PRIOR FILING DATE: 1997-05-05
; NUMBER OF SEQ ID NOS: 129
; SEQ ID NO 2
; LENGTH: 931
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-039-2

Query Match 86.2%; Score 496.6; DB 10; Length 931;
Best Local Similarity 91.5%; Pred. No. 1.1e-148;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATCAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
DB 99 ATGAAAGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 158
QY 61 TTACGAAGGCTTCGATGATATTCGCAAAATCTCTCTAAGAAGAGTGGGAAAAGATG 120

DB 159 ATCCAAAAGGCTTCGATGATATTCGCAAAATCTCTCTAAGAAGAGTGGGAAAAGATG 218
QY 121 AAATCCTCGAGAAAATCGTCTATGTATATGAAGCTAAACTATGAGGTCTAGCTAAA 180
DB 219 AAAGCCTCGAGAAAATCTTCTATGTATATGAAGAAAGTATGAGGCTATGACTAAA 278
QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTCCTATGCTAGTAAACCGGCTGCGAGCTTCCAC 240
DB 279 CTAGGTTTCAAGGCCACCTCCACCTTCCTATGCTAGTAAACCGGCTGCGAGCTTCCAG 338
QY 241 GGAATGATTTTGGTAAACGATCGAAACACACAGGAATCAGGTTTGAACCTCTCAGATGACT 300
DB 339 GGAATGATTTGGTAAACGATCGAAACACACAGGAATCAGGTTTGAACCTCTCAGATGACT 398
QY 301 TTGGGAGCCTCCAGAAATCTTCCGAAAGTATGCGCAAGTATGCGCAAGCCAGAGAGAA 360
DB 399 TTGGGAGCCTCCAGAAATCTTCCGAAAGTATGCGCAAGTATGCGCAAGCCAGAGAGAA 456
QY 361 AATGTTTGAAGGAGTGCAGAGGATCTGCGCAAGTATGCGCAAGTATGCGCAAGCCAGAGAG 420
DB 459 AATGATTTGGGAGGAGTGCAGAGGATCTGCGCAAGTATGCGCAAGTATGCGCAAGCCAGAGAG 518
QY 421 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTTAAACAGACATCTGGACCCCAAAAGGGGG 480
DB 519 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTTAAACAGACATCTGGACCCCAAAAGGGGG 578
QY 481 AAACATGCTGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATGAGAGATC 540
DB 579 GAAATGCTGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATGAGAGATC 638
QY 541 AGCGACCTTGAGGAAGATGACGAGTAACCTCCCTC 575
DB 639 AGCGACCTTGAGGAAGATGACGAGTAACCTCCCTC 673

RESULT 9

US-10-177-277-2
; Sequence 2, Application US/10177277
; Publication No. US20030185844A1
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; APPLICANT: Ramensee, Hans Georg
; APPLICANT: Stevanovic, Stefan
; TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Determining Presence of an SSX Gene, Peptides Derived From Said SSX Gene
; TITLE OF INVENTION: Expression of an SSX Gene, and Uses Thereof
; FILE REFERENCE: LUD 5556.1
; CURRENT APPLICATION NUMBER: US/10/177,277
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US/09/344,040
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US/09/105,839
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US/08/851,130
; PRIOR FILING DATE: 1997-05-05
; NUMBER OF SEQ ID NOS: 132
; SEQ ID NO 2
; LENGTH: 931
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-277-2

Query Match 86.2%; Score 496.6; DB 14; Length 931;
Best Local Similarity 91.5%; Pred. No. 1.1e-148;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ATCAACGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
DB 99 ATGAAAGGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 158

QY 61 TTACGAAAGCCCTTCGATGATATGCGAAATACCTCTCTAAGAAAGAGTGGAAAGATG 120
 Db 159 ATCCAAAGCCCTTCGATGATATGCGAAATACCTCTCTAAGAAAGAGTGGAAAGATG 218
 QY 121 AAATCCTCGAGAAATCGTCTATGATATGCGAAATACCTCTCTAAGAAAGAGTGGAAAGATG 180
 Db 219 AAGCCCTCGAGAAATCGTCTATGATATGCGAAATACCTCTCTAAGAAAGAGTGGAAAGATG 278
 QY 181 CTAGGTTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 240
 Db 279 CTAGGTTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 338
 QY 241 GGAATGATTTGGTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 300
 Db 339 GGAATGATTTGGTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 398
 QY 301 TTCCGAGCCCTCCAGAAATCTTCCGAGAGATCATGCCCAAGAGTCAAGTCAAGTCAAGT 360
 Db 399 TTCCGAGCCCTCCAGAAATCTTCCGAGAGATCATGCCCAAGAGTCAAGTCAAGTCAAGT 458
 QY 361 AATGGTTTGAAGAGTCCAGAGGATCTGCCCAAGAGTCAAGTCAAGTCAAGTCAAGTCAAGT 420
 Db 459 AATGATTCGAGAGTCCAGAGGATCTGCCCAAGAGTCAAGTCAAGTCAAGTCAAGTCAAGT 518
 QY 421 CCCCAGGAAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 480
 Db 519 CCCCAGGAAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 578
 QY 481 AATGATTCGAGAGTCCAGAGGATCTGCCCAAGAGTCAAGTCAAGTCAAGTCAAGTCAAGT 540
 Db 579 AATGATTCGAGAGTCCAGAGGATCTGCCCAAGAGTCAAGTCAAGTCAAGTCAAGTCAAGT 638
 QY 541 AGCGACCTCGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 575
 Db 639 AGCGACCTCGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 673

RESULT 10

US-09-954-531-1000
 ; Sequence 1000, Application US/09954531
 ; Patent No. US20020165180A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Weaver, Zoe
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
 ; FILE REFERENCE: 689290-77
 ; CURRENT APPLICATION NUMBER: US/09/954,531
 ; CURRENT FILING DATE: 2002-05-02
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: US/60/233,133
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: US/60/234,009
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US/60/234,034
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US/60/234,509
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US/60/234,567
 ; NUMBER OF SEQ ID NOS: 1392
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 1000
 ; LENGTH: 1309
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1309)
 ; OTHER INFORMATION: n=a,t,g or c
 US-09-954-531-1000

Query Match 86.2%; Score 496.6; DB 9; Length 1309;
 Best Local Similarity 91.5%; Pred. No. 1.4e-148;
 Matches 526; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATCTCAATATCAGAGAAG 60
 Db 99 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATCTCAATATCAGAGAAG 158
 QY 61 TTACGAAGGCTTCGATGATATGCGAAATACCTCTCTAAGAAAGAGTGGAAAGATG 120
 Db 159 ATCCAAAGCCCTTCGATGATATGCGAAATACCTCTCTAAGAAAGAGTGGAAAGATG 218
 QY 121 AAATCCTCGAGAAATCGTCTATGATATGCGAAATACCTCTCTAAGAAAGAGTGGAAAGATG 180
 Db 219 AAGCCCTCGAGAAATCGTCTATGATATGCGAAATACCTCTCTAAGAAAGAGTGGAAAGATG 278
 QY 181 CTAGGTTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 240
 Db 279 CTAGGTTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 338
 QY 241 GGAATGATTTGGTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 300
 Db 339 GGAATGATTTGGTAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 398
 QY 301 TTCCGAGCCCTCCAGAAATCTTCCGAGAGATCATGCCCAAGAGTCAAGTCAAGTCAAGT 360
 Db 399 TTCCGAGCCCTCCAGAAATCTTCCGAGAGATCATGCCCAAGAGTCAAGTCAAGTCAAGT 458
 QY 361 AATGGTTTGAAGAGTCCAGAGGATCTGCCCAAGAGTCAAGTCAAGTCAAGTCAAGTCAAGT 420
 Db 459 AATGATTCGAGAGTCCAGAGGATCTGCCCAAGAGTCAAGTCAAGTCAAGTCAAGTCAAGT 518
 QY 421 CCCCAGGAAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 480
 Db 519 CCCCAGGAAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 578
 QY 481 AATGATTCGAGAGTCCAGAGGATCTGCCCAAGAGTCAAGTCAAGTCAAGTCAAGTCAAGT 540
 Db 579 AATGATTCGAGAGTCCAGAGGATCTGCCCAAGAGTCAAGTCAAGTCAAGTCAAGTCAAGT 638
 QY 541 AGCGACCTCGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 575
 Db 639 AGCGACCTCGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 673

RESULT 11

US-09-954-531-1392
 ; Sequence 1392, Application US/09954531
 ; Patent No. US20020165180A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Weaver, Zoe
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Ca
 ; FILE REFERENCE: 689290-77
 ; CURRENT APPLICATION NUMBER: US/09/954,531
 ; CURRENT FILING DATE: 2002-05-02
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: US/60/233,133
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: US/60/234,009
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US/60/234,034
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US/60/234,509
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: US/60/234,567
 ; NUMBER OF SEQ ID NOS: 1392
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 1392
 ; LENGTH: 1309
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(1309)
 ; OTHER INFORMATION: n=a,t,g or c

ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. US20030023057Alman D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5480
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3168
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 576 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-975-856-2

Query Match 83.7%; Score 482.2; DB 10; Length 576;
Best Local Similarity 89.9%; Pred. No. 3.7e-144;
Matches 517; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCGCTTTCGAAGGAGACCCAGGGATGATGCTCAATATCAGAGAG 60
Db 1 ATGAACGGAGACGACGCGCTTTCGAAGGAGACCCAGGGATGATGCTCAATATCAGAGAG 60

QY 61 TTACGAAAGCGCTTCGATGATATGCCAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
Db 61 ATGCAAAAGCGCTTCGATGATATGCCAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120

QY 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCAATGCTAA 180
Db 121 AAAGCTCGGAGAAATCATCTATGTATATGAAGAAAGATGAGGCCATGACTAA 180

QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTCCAC 240
Db 181 CTAGGTTTCAAGGCGACCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTCCAC 240

QY 241 GGAATGATTTTGGTAAAGATCGAAACACAGGAATCAGTTGAAGCTCTCAGATGACT 300
Db 241 GGAATGATTTTGGTAAAGATCGAAACACAGGAATCAGTTGAAGCTCTCAGATGACT 300

QY 301 TTCCGAGAGCTCCAGAGATCTTCCGAGAGATCATGCCAAGAGCCAGAGAGAGAA 360
Db 301 TTCCGAGAGCTCCAGAGATCTTCCGAGAGATCATGCCGAGAGCCAGAGAGAGAA 360

QY 361 AATGTTTGAAGAAAGTCCAGAGGATCTGCGCCACAAAATGATGGGAAACAGCTGTGC 420
Db 361 AATGATTCAAAGGGAGTGCAGAGCATCTGCGCCACAGAAACAATGGGAAACAGCTGCGC 420

QY 421 CCCCAGGAAATCCCAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAGGGGG 480
Db 421 CCCTCAGGAAATCAATACCTTGGAGAGATTAAACAGACATCTGGACCCAAAGGGGG 480

QY 481 AAACATGCTGGACCCACAGATCTGCGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
Db 481 AAACATGCTGGACCCACAGATCTGCGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540

QY 541 AGCGACCTGAGAGATGACGAGTAACCTCCCTC 575
Db 541 AGCGACCTGCGGAAGATGACGAGTAACCTCCCTC 575

RESULT 14
US-09-833-039-6
; Sequence 6, Application US/09833039
; Publication No. US20030175960A1
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; TITLE OF INVENTION: Tumor Associated Peptide and Uses Thereof
; FILE REFERENCE: LUD 5622.1
; CURRENT APPLICATION NUMBER: US/09/833,039
; CURRENT FILING DATE: 2001-04-12

; PRIOR APPLICATION NUMBER: US 09/409,455
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: US 09/344,040
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/105,839
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US 08/851,130
; PRIOR FILING DATE: 1997-05-05
; NUMBER OF SEQ ID NOS: 129
; SEQ ID NO 6
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-039-6

Query Match 83.7%; Score 482.2; DB 10; Length 576;
Best Local Similarity 89.9%; Pred. No. 3.7e-144;
Matches 517; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACGCGCTTTCGAAGGAGACCCAGGGATGATGCTCAATATCAGAGAG 60
Db 1 ATGAACGGAGACGACGCGCTTTCGAAGGAGACCCAGGGATGATGCTCAATATCAGAGAG 60

QY 61 TTACGAAAGCGCTTCGATGATATGCCAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120
Db 61 ATGCAAAAGCGCTTCGATGATATGCCAAATATCTTCTTAAGAAAGAGTGGGAAAGATG 120

QY 121 AAATCTCGGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCAATGCTAA 180
Db 121 AAAGCTCGGAGAAATCATCTATGTATATGAAGAAAGATGAGGCCATGACTAA 180

QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTTCCAC 240
Db 181 CTAGGTTTCAAGGCGACCTCCACCTTTCATGCGTAGTAAACGGGCTGCAGACTTTCCAC 240

QY 241 GGAATGATTTTGGTAAAGATCGAAACACAGGAATCAGTTGAAGCTCTCAGATGACT 300
Db 241 GGAATGATTTTGGTAAAGATCGAAACACAGGAATCAGTTGAAGCTCTCAGATGACT 300

QY 301 TTCCGAGAGCTCCAGAGATCTTCCGAGAGATCATGCCAAGAGCCAGAGAGAGAA 360
Db 301 TTCCGAGAGCTCCAGAGATCTTCCGAGAGATCATGCCGAGAGCCAGAGAGAGAA 360

QY 361 AATGTTTGAAGAAAGTCCAGAGGATCTGCGCCACAAAATGATGGGAAACAGCTGTGC 420
Db 361 AATGATTCAAAGGGAGTGCAGAGCATCTGCGCCACAGAAACAATGGGAAACAGCTGCGC 420

QY 421 CCCCAGGAAATCCCAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCAAAGGGGG 480
Db 421 CCCTCAGGAAATCAATACCTTGGAGAGATTAAACAGACATCTGGACCCAAAGGGGG 480

QY 481 AAACATGCTGGACCCACAGATCTGCGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540
Db 481 AAACATGCTGGACCCACAGATCTGCGTGAGAGAAAGCAGCTGGTGTATGAAGAGATC 540

QY 541 AGCGACCTGAGAGATGACGAGTAACCTCCCTC 575
Db 541 AGCGACCTGCGGAAGATGACGAGTAACCTCCCTC 575

RESULT 15
US-10-177-277-6
; Sequence 6, Application US/10177277
; Publication No. US20030185844A1
; GENERAL INFORMATION:
; APPLICANT: Tureci, Ozlem
; APPLICANT: Sahin, Ugur
; APPLICANT: Pfreundschuh, Michael
; APPLICANT: Ramensee, Hans Georg
; APPLICANT: Stevanovic, Stefan
; TITLE OF INVENTION: Method for Determining Presence of Cancer in a Sample By Detect
; TITLE OF INVENTION: Expression of an SSX Gene, Peptides Derived From Said SSX Gene
; TITLE OF INVENTION: Gene, and Uses Thereof


```
; FILE REFERENCE: IUD 5556.1
; CURRENT APPLICATION NUMBER: US/10/177,277
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US/09/344,040
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US/09/105,839
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: US/08/851,130
; PRIOR FILING DATE: 1997-05-05
; NUMBER OF SEQ ID NOS: 132
; SEQ ID NO 6
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-277-6

Query Match      83.7%; Score 482.2; DB 14; Length 576;
Best Local Similarity 89.9%; Pred. No. 3.7e-144;
Matches 517; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY      1 ATGAACGGAGACGACGCCCTTTCAGAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
Db      |||
QY      1 ATGAACGGAGACGACGCCCTTTCAGAGGACCTAGGGTTGTTCTCAATATCAGAGAAG 60
Db      |||
QY      61 TTAGAAAGCCCTTCGATGATATTCGCAAAATCTTCTTAAGAAAGATGGGAAAAGATG 120
Db      |||
QY      61 ATGCAAAAGCCCTTCGATGATATTCGCAAAATCTTCTCTGAGAAAGATGGGAAAAGATG 120
Db      |||
QY      121 AAATCCTCGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGTCTGACTAAA 180
Db      |||
QY      121 AAAGCCTCGAGAAATCTCTATGTATATGAAGAAAGATATGAGGCCATGACTAAA 180
Db      |||
QY      181 CTAGGTTTCAAGTCCACCTCCACCTTTTCATGCGTAGTAAACGGGCTGACACTTCCAC 240
Db      |||
QY      181 CTAGGTTTCAAGGCCACCTCCACCTTTTCATGCGTAAATAACGGGTGCGACACTTCCAG 240
Db      |||
QY      241 GGAATGATTTTGGTAAAGTCAAGTCAAGACACAGGATCAGGTTGAGCGTCTCAGATGACT 300
Db      |||
QY      241 GGAATGATTTTGAATGACCTTAACCTGAGGATCAGGTTGAGGATCAGGTTGAGGATCAGT 300
Db      |||
QY      301 TTCGGCAGGCTCCAGAGAATCTTCCGAAAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360
Db      |||
QY      301 TTCGGCAGGCTCCAGGGAATCTTCCGAAAGATCAGCCCGAGAGCCAGCAGAGGAAGGA 360
Db      |||
QY      361 AATGTTTGAAGAAAGTGCAGAGGATCTGCCCAACAAATGATCGGAAACAGCTGTGC 420
Db      |||
QY      361 AATGATTCAAAGGGAGTGCAGAGGATCTGCCCAACAAATGATCGGAAACAGCTGTGC 420
Db      |||
QY      421 CCCCCGGAAATCCAAAGTACCTTGGAGAGATTAAACAGACATCTGGACCCCAAAAGGGG 480
Db      |||
QY      421 CCTCAGGAAACTAATACCTCTGAGAGGTTTACAGACATCTGGACCCCAAAAGGGG 480
Db      |||
QY      481 AAACATGCTGACCCACAGACTGCTGAGAGAAAGCAGCTGGTGTATTAAGAGATC 540
Db      |||
QY      481 AAACATGCTGACCCACAGAGTGTGAGAGAAAGCAACTGGTGTATTAAGAGATC 540
Db      |||
QY      541 AGCGACCTTGAGGAAGATGACGAGTAATCTCCCTC 575
Db      |||
QY      541 AGCGACCTTGCGGAAGATGACGAGTAATCTCCCTC 575
Db      |||
```

Search completed: March 31, 2004, 18:32:55
Job time : 402 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 31, 2004, 13:39:56 ; Search time 14 Seconds
(without alignments)
4284.630 Million cell updates/sec

Title: US-09-975-856-1

Perfect score: 1025

Sequence: 1 ATGACGAGACGACGCGCTT.....ATGACGAGTAACCTCCCTCG 576

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgm2.1/USPTO.spool.p/US09975856/runat_31032004_132903_14738/app.query.fasta_1.775
-DB=SwissProt_42 -OPMT=fastan -SUFFIX=xsp -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09975856.cgm 1 16 @runat_31032004_132903_14738 -NCPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1004	98.0	188	1 SX4 HUMAN	O60224 homo sapien
2	801	78.1	188	1 SX3 HUMAN	Q99909 homo sapien
3	798	77.9	188	1 SX2 HUMAN	Q16385 homo sapien
4	784	76.5	188	1 SX1 HUMAN	Q16384 homo sapien
5	780	76.1	188	1 SX5 HUMAN	O60225 homo sapien
6	113.5	11.1	289	1 ZN75 HUMAN	P51815 homo sapien
7	101	9.9	642	1 ZN75 HUMAN	Q8td17 homo sapien
8	94	9.2	595	1 Z317 HUMAN	Q96pg6 homo sapien
C 9	89.5	8.9	1505	1 PK3G RAT	O70173 rattus norv
C 10	88.5	8.8	531	1 FMO1 CANFA	Q95la2 canis fami
11	88	8.6	743	1 REP1 MOUSE	Q54916 mus musculu
C 12	87.5	8.7	531	1 FMO1 PIG	P16549 sus scrofa
13	87	8.5	744	1 REP1 HUMAN	Q96d71 homo sapien
14	86.5	8.4	396	1 SX11 CHICK	P48435 Gallus gall
15	86.5	8.4	446	1 ZN38 HUMAN	P17036 homo sapien
16	86.5	8.4	488	1 ZP92 MOUSE	Q62396 mus musculu
17	86.5	8.4	572	1 TC17 MOUSE	Q61751 mus musculu
18	86	8.4	480	1 WR61 ARATH	Q8vww6 arabidopsis

19	86	8.4	834	1 CASL HUMAN	Q14511 homo sapien
20	85.5	8.3	680	1 Z334 HUMAN	Q9hcz1 homo sapien
21	85.5	8.3	738	1 ZN84 HUMAN	P51523 homo sapien
22	84.5	8.2	429	1 HIXX METJA	Q58851 methanococc
23	84.5	8.2	463	1 Z331 HUMAN	Q9rxk6 homo sapien
C 24	84.5	8.4	534	1 FMO1 RABIT	P17636 oryctolagus
25	84	8.2	524	1 Z479 HUMAN	Q6je14 homo sapien
26	83.5	8.1	636	1 ZF90 MOUSE	Q61967 mus musculu
27	83.5	8.1	1102	1 ANP HUMAN	Q9H2D0 homo sapien
28	83.5	8.1	1630	1 MSF1 FLAFW	P04932 plasmodium
29	83.5	8.1	1639	1 MSF1 FLAFW	P04933 plasmodium
30	82.5	8.0	553	1 Z324 HUMAN	O75467 homo sapien
31	82.5	8.0	643	1 ZN74 HUMAN	P16587 beetle wester
32	82.5	8.0	1035	1 RRPO BWYVE	P09507 beetle wester
33	81.5	8.0	504	1 Z205 HUMAN	O95201 homo sapien
34	81.5	8.0	639	1 GLCX SOYEN	P11827 glycine max
35	81.5	8.0	1395	1 SF41 YEAST	P38904 saccharomyc
36	81.5	8.0	1443	1 DPO3 MYCPN	P75080 mycoplasma
37	81	7.9	458	1 ZN19 HUMAN	P17023 homo sapien
38	81	7.9	751	1 Z337 HUMAN	Q9Y3M9 homo sapien
C 39	81	8.0	933	1 PERT CANFA	Q8hyb7 canis fami
40	80.5	7.9	439	1 ZN10 HUMAN	P21506 homo sapien
41	80.5	7.9	470	1 Z436 HUMAN	Q9c0f3 homo sapien
42	80.5	7.9	659	1 Z304 HUMAN	Q9hcx3 homo sapien
43	80.5	7.9	998	1 GFPI HUMAN	P78347 h general t
C 44	79.5	7.9	531	1 FMO1 HUMAN	Q01740 homo sapien
45	79.5	7.8	626	1 Z471 HUMAN	Q9bxb2 homo sapien

ALIGNMENTS

RESULT 1

SSX4_HUMAN
ID SSX4_HUMAN STANDARD; PRT; 188 AA.
AC O60224; O9UTJ9;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SSX4 protein.
GN SSX4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98021352; PubMed=9378559;
RA Gure A.O., Tuereci O., Sahin U., Tsang S., Scanlan M.J., Jager E.,
RA Knuth A., Pfrendschuh M., Old L.J., Chen Y.-T.;
RA "SSX: a multigene family with several members transcribed in normal
RT testis and human cancer.";
RL Int. J. Cancer 72:965-971(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Urinary bladder;
RX MEDLINE=22388257; PubMed=12477932;
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsien F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Schnerch A., Schein J.E., Jones S.O.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length
 RL human and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP SEQUENCE OF 1-155 FROM N.A.
 RA Blechschmidt K., Nyakatura G., Strom T.M., Drescher B., Menzel U.,
 RA Meindl A., Rosenthal A.;
 .RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Could act as a modulator of transcription.
 CC -!- SIMILARITY: Belongs to the SSX family.
 CC -!- SIMILARITY: Contains 1 KRAB-related domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U0841; AAC05820.1; .
 DR EMBL; BC005325; AAH05325.1; .
 DR EMBL; AF196772; AAF06796.1; .
 DR Genew; HGNC:11338; SSX4.
 DR MIM; 300326; .
 DR InterPro; IPR001909; KRAB.
 DR InterPro; IPR003655; KRAB_related.
 DR Pfam; PF01352; KRAB; 1.
 DR SMART; SM00349; KRAB; 1.
 DR PROSITE; PS0806; KRAB_RELATED; 1.
 KW Multigene family; Transcription regulation.
 FT DOMAIN 20 83
 SQ SEQUENCE 188 AA; 21858 MW; 4ACA2A8737507AE5 CRC64;

Alignment Scores:

Pred. No.: 1.37e-91 Length: 188
 Score: 1004.00 Matches: 188
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 97.95% Indels: 0
 DB: 1 Gaps: 0

US-09-975-856-1 (1-576) x SSX4_HUMAN (1-188)

QY 1 ATGACGGAGAGAGCGCTTTCAGAGGAGACCCAGGAGATGCTCAATATCAGAGAG 60
 Db 1 MetAsnGlyAspAlaPheAlaArgProArgAspAlaGlnSerGluLys 20
 QY 61 TTACGAAAGCGCTTCGATGATATCCCAATACCTTCTTAAGAAAGAGTGGGAAAAGATG 120
 Db 21 LeuArgLysAlaPheAspAlaLysTyrPheSerLysLysGluTyrGluLysMet 40
 QY 121 AAATCTCGGAGAAATCGTCTATGCTATATGAGCTAAACTATGAGTCACTAA 180
 Db 41 LysSerSerGluLysValTyrValTyrMetLysLeuAsnTyrGluValMetThrLys 60
 QY 181 CTAGTTTCAAGGTCAACCTCCACCTTTCATGCTAGTAAACGGCTGCAGACTTCCAC 240
 Db 61 LeuGlyPheLysValThrLeuProPheMetArgSerLysArgAlaAspPheHis 80
 QY 241 GGGATGATTTTGGTAACGATCGAAACACAGCAATCAGTTTGAACGCTCCAGTACT 300
 Db 81 GlyAsnAspPheGlyAsnAspArgAsnHisArgAsnGlnValGluArgProGlnMetThr 100
 QY 301 TTCGGCAGCTCCAGAGATCTTCCGAGATCATGCCAAGATCATGCCAAGACCCAGCAGAGAA 360
 Db 101 PheGlySerLeuGlnArgIlePheProLysIleMetProLysLysProAlaGluGlu 120
 QY 361 AATGTTTGAAGAGATGCGCAGGAGCTTCCGCCACAAATGATGGAAACAGCTGTC 420
 Db 121 AsnGlyLeuLysGluValProGluAlaSerGlyProGlnAsnAspGlyLysGlnLeu 140
 QY 421 CCCCAGGGAATCCCAAGTACCTTGGGAGAGATTAACAAGACATCTGGACCCCAAGGGGG 480

Db 141 ProProGlyAsnProSerThrLeuGluLysIleAsnLysThrSerGlyProLysArgGly 160
 QY 481 AAACATGCTGAGACCCACACAGCTGCTGAGAGAAAGACGCTGCTGTTATGAAGATC 540
 Db 161 LysHisAlaTyrThrHisArgLeuArgGluAArgLysGlnLeuValValTyrGluGlu 180
 QY 541 AGCGACCTGAGAGAGATGACGAG 564
 Db 181 SerAspProGluGluAspAspGlu 188

RESULT 2

SSX3_HUMAN
 ID SSX3_HUMAN STANDARD; PRT; 188 AA.
 AC OS9979; O60223;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE SSX3 protein.
 GN SSX3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fibrosarcoma, and Testis;
 RX MEDLINE=96302330; PubMed=8697803;
 RA de Leeuw B., Balemans M., Geurts van Kessel A.;
 RT "A novel Kruppel-associated box containing the SSX gene (SSX3) on the
 RT human X chromosome is not implicated in t(X;18)-positive synovial
 RT sarcomas.";
 RL Cytogenet. Cell Genet. 73:179-183 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98021352; PubMed=9378559;
 RA Gure A.O., Tuereci O., Sahin U., Tsang S., Scanlan M.J., Jager E.,
 RA Kruth A., Pfeundscher M., Old L.J., Chen Y.-T.;
 RT "SSX: a multigene family with several members transcribed in normal
 RT testis and human cancer.";
 RL Int. J. Cancer 72:965-971 (1997).
 CC -!- FUNCTION: Could act as a modulator of transcription.
 CC -!- SIMILARITY: Belongs to the SSX family.
 CC -!- SIMILARITY: Contains 1 KRAB-related domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; S82471; AAB37436.2; .
 DR EMBL; U90840; AAC05819.1; .
 DR Genew; HGNC:11337; SSX3.
 DR MIM; 300325; .
 DR InterPro; IPR001909; KRAB.
 DR InterPro; IPR003655; KRAB_related.
 DR Pfam; PF01352; KRAB; 1.
 DR SMART; SM00349; KRAB; 1.
 DR PROSITE; PS0806; KRAB_RELATED; 1.
 KW Multigene family; Transcription regulation.
 FT DOMAIN 20 83
 FT CONFLICT 95 95 L -> Q (IN REF. 2).
 SQ SEQUENCE 188 AA; 21682 MW; 368C7487C1450E72 CRC64;

Alignment Scores:

Pred. No.: 1.85e-71 Length: 188
 Score: 801.00 Matches: 152
 Percent Similarity: 84.57% Conservative: 7
 Best Local Similarity: 80.85% Mismatches: 29
 Query Match: 78.15% Indels: 0

```
DB: 1 1 Gaps: 0
US-09-975-856-1 (1-576) x SSX2_HUMAN (1-188)
QY 1 ATGAACGGAGACACCCCTTCGACGAGACCCAGGATGATGCTCAATATCAGAGAG 60
Db 1 McAsnGlyAspAspThrPheAlaArgAsgProThrValGlyAlaGlnIleProGluLys 20
QY 61 TTACGAAGAGCCCTTCGATGATATGTCACAAATACTTCTTAAGAAAGAGTGGAAAGATG 120
Db 21 IleGlnLysAlaPheAspAspIleAlaLysTyPheSerLysGlnGluTrpGluLysMet 40
QY 121 AATCTCCGAGAAATCGTCTATGCTATATGAAGCTAACTAAGTATGAGTCTAGCTATAA 180
Db 41 LysValSerGluLysIleValTyMetLysArgLysTyGluAlaMetThrLys 60
QY 181 CTAGGTTTCAAGGTCACCTCCACCTTCATCGCTAGTAAACGGGCTCGACACTCCAC 240
Db 61 LeuGlyPheLysAlaIleLeuProSerPheMetArgAsnLysArgValThrAspPheGln 80
QY 241 GGAATGATTTGGTAACGATCGAACACACAGGAATCAGGTGTAACGCTCCTCAGATGACT 300
Db 81 GlyAsnAspPheAspAsnAspProAsnArgGlyAsnGlnValLeuArgProGlnMetThr 100
QY 301 TTCCGGCAGCCTCCAGAGAAATCTTCCGAGAGATCATGCCAAGAGCCAGCAGAGAGAA 360
Db 101 PheGlyArgLeuGlnGlyIlePheProLysIleMetProLysLysProIleGluGluGly 120
QY 361 ATGTTGTTGAAGAAATGTCAGAGGCAATTCGCCCAACAAATATGATGGGAAACAGCTGTC 420
Db 121 AsnValSerLysGluValProGluAlaSerGlyProGlnAsnAspGlyLysGlnLeuLys 140
QY 421 CCCCAGGAAATCCCAAGTACCTTGGAGAGAGATTAACAAGACATCTGCACCCAAAGGGGG 480
Db 141 ProProGlyLysProThrThrSerGluLysIleAsnMetIleSerGlyProLysArgGly 160
QY 481 AAACATGCTCGGACCCACAGATGCTGTCGAGAGAAACAGCTGCTGTATGAAGATC 540
Db 161 GluHisAlaIleThrHisArgLeuArgGluArgGluArgLysGlnLeuValIleTyGluGluIle 180
QY 541 ACGGACCCCTCAGAGAGATGACGAG 564
Db 181 SerAspProGluGluAspGlu 188

RESULT 3
SSX2_HUMAN STANDARD; PRT; 188 AA.
ID Q16351; Q16404; Q96IP7;
AC 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE SSX2 protein (synovial sarcoma, X breakpoint 2) (SSX) (HOM-MEL-40).
GN SSX2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RP TISSUE=Testis;
RX MEDLINE=95292974; PubMed=7539744;
RA Crew A.J., Clark J., Fisher C., Gill S., Grimer R., Chand A.,
RA Shipley J., Gusterson B.A., Cooper C.S.;
RT "Fusion of SYT to two genes, SSX1 and SSX2, encoding proteins with
RT homology to the Kruppel-associated box in human synovial sarcoma.";
RL EMBL J. 14:2333-2340(1995).
[2]
SEQUENCE FROM N.A.
RP TISSUE=Skin;
RX MEDLINE=22368257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Tqulliano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Guarnare F.H.,
RA Villalón D.K., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.B.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[3]
SEQUENCE OF 111-188 FROM N.A.
RP TISSUE=Synovial sarcoma;
RX MEDLINE=95384157; PubMed=7655467;
RA de Leeuw B., Bailemans M., Olde Weghuis D., Geurts van Kessel A.;
RT "Identification of two alternative fusion genes, SYT-SSX1 and
RT SYT-SSX2, in t(X;18)(p11.2;q11.2)-positive synovial sarcomas.";
RL Hum. Mol. Genet. 4:1097-1099(1995).
[4]
SEQUENCE OF 111-188 FROM N.A. (SSX2-SSX2 FUSION PROTEIN).
RP TISSUE=Synovial sarcoma;
RX MEDLINE=95038836; PubMed=7951320;
RA Clark J., Roques P.J., Crew A.J., Gill S., Shipley J., Chan A.M.-L.,
RA Gusterson B.A., Cooper C.S.;
RT "Identification of novel genes, SYT and SSX, involved in the
RT t(X;18)(p11.2;q11.2) translocation found in human synovial sarcoma.";
RL Nat. Genet. 7:502-508(1994).
[5]
SEQUENCE OF 68-116 FROM N.A. (SSX2-SSX2 FUSION PROTEIN).
RX MEDLINE=96094743; PubMed=7495284;
RA Fligman I., Lonardo F., Jhanwar S.C., Gerald W.L., Woodruff J.,
RA Ladanyi M.;
RT "Molecular diagnosis of synovial sarcoma and characterization of a
RT variant SYT-SSX2 fusion transcript.";
RL Am. J. Pathol. 147:1592-1599(1995).
CC -!- FUNCTION: Could act as a modulator of transcription.
CC -!- TISSUE SPECIFICITY: Expressed at high level in the testis.
CC Expressed at low level in thyroid. Not detected in tonsil, colon,
CC lung, spleen, prostate, kidney, striated and smooth muscles.
CC Detected in rhabdomyosarcoma and fibrosarcoma cell lines. Not
CC detected in mesenchymal and epithelial cell lines.
CC -!- DISEASE: Participates in a t(X;18)(p11.2;q11.2) chromosomal
CC translocation that is specifically found in more than 80% of
CC synovial sarcoma and produces the SSX2-SSX1 or SSX2-SSX2 fusion
CC products. These hybrid proteins are probably responsible for
CC transforming activity. Heterogeneity in the position of the
CC breakpoint can occur (low frequency).
CC -!- SIMILARITY: Belongs to the SSX family.
CC -!- SIMILARITY: Contains 1 KRAB-related domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X86175; CAA60111.1; --
DR EMBL; BC007343; AA07343.1; --
DR EMBL; BC016957; AA016957.1; --
DR EMBL; S79332; AAB35379.1; --
DR EMBL; X79200; --; NOT_ANNOTATED_CDS.
DR EMBL; S79894; AAB35674.1; ALT_INIT.
DR PIR; S55058; S55058.
DR Genew; HGNC:11336; SSX2.
```

```
DR MTM; 300192; ..
DR GO; 0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003655; KRAB_related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SMO0349; KRAB; 1.
DR PROSITE; PSS0806; KRAB RELATED; 1.
KW Chromosomal translocation; Proto-oncogene; Multigene family;
KW Transcription regulation.
FT DOMAIN 20 83 KRAB-RELATED.
FT SITE 68 69 BREAKPOINT FOR TRANSLOCATION TO FORM THE
FT SITE 110 111 SSX1-SSX2 FUSION PROTEIN (RARE).
FT SITE 169 169 BREAKPOINT FOR TRANSLOCATION TO FORM THE
FT CONFLICT 169 169 SSX1-SSX2 FUSION PROTEIN.
FT SEQUENCE 188 AA; 21620 MW; BF5D18AA5F45B1B1 CRC64;
Alignment Scores:
Pred. No.: 3.66e-71 Length: 188
Score: 798.00 Matches: 150
Percent Similarity: 86.17% Conservative: 12
Best Local Similarity: 79.79% Mismatches: 26
Query Match: 77.85% Indels: 0
DB: 1 Gaps: 0
US-09-975-856-1 (1-576) x SSX2_HUMAN (1-188)
QY 1 ATGAACGAGACGACGCGCTTTGCAAGGAGACCCAGGAGATGCTCAATATATCAGAGAAG 60
DQ 1 MetAsnGlyAspAlaPheAlaArgProThrValGlyAlaGlnIleProGluLys 20
QY 61 TTACGAAGCCCTCGATGATATTCCTCAATATATTCCTCAAGAAAGTGGGAAGATG 120
DQ 21 IleGlnLysAlaPheAspAlaPheAlaLysPheSerLysGluLysMet 40
QY 121 AAATCCCTCGAGAAATCGCTATGTATATGAAGTAACTATGAGTGCATGACTAA 180
DQ 41 LysAlaSerGluLysIlePheTyrValTyrMetLysArgLysTyrGluAlaMetThrLys 60
QY 181 CTAGTTTCAAGTCCACCTCCACCTTCATCGGTAGTAAAGCGGTGCAGACTTCCAC 240
DQ 61 LeuGlyPheLysAlaThrLeuProProPheMetCysAsnLysArgAlaGluAspPheGln 80
QY 241 GGGATATTTGTTGATACGATCGAAGACACAGGATCAGTCTGAACCTCTCAGATGACT 300
DQ 81 GlyAsnAspLeuAspAsnAspProAsnArgGlyAsnGlnValGluArgProGlnMetThr 100
QY 301 TTGGCAGCTCCAGAGATCTTCCGAGATCATGCCAAGATCCAGAGAGAGAGAA 360
DQ 101 PheGlyArgLeuGlnGlyIleSerProLysIleMetProLysLysProAlaGluGluGly 120
QY 361 AATGTTTGAGGAGTGCAGAGGATCTGGCCCAAAATGATGGAAACAGCTGTGC 420
DQ 121 AsnAspSerGluGluValProGluAlaSerGlyProGlnAsnAspGlyLysGluLeuLys 140
QY 421 CCCCCGGGAAATCCAGTACCTTGGGAGAGATTAAACAGATCTGGACCCCAAGGGGG 480
DQ 141 ProProGlyLysProThrThrSerGluLysIleHisGluArgSerGlyProLysArgGly 160
QY 481 AAATCCCTCGAGAGATCTCGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 540
DQ 161 GluHisAlaThrThrHisArgLeuArgGluArgLysGlnLeuValIleTyrGluGluIle 180
QY 541 AGCGACCTCGAGAGATGACGAG 564
DQ 181 SerAspProGluGluAspAspGlu 188
RESULT 4
SSX1_HUMAN STANDARD; PRT; 188 AA.
AC Q16384;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
```

10-OCT-2003 (Rel. 42, Last annotation update)
SSX1 protein (Synovial sarcoma, X breakpoint 1).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibrosarcoma;
RX MEDLINE=95292974; PubMed=7539744;
RA Crew A.J., Clark J., Fisher C., Gill S., Grimer R., Chand A.,
RA Shipley J., Gusteron B.A., Cooper C.S.;
RT "Fusion of SYT to two genes, SSX1 and SSX2, encoding proteins with
RT homology to the Kruppel-associated box in human synovial sarcoma.";
RL EMBO J. 14:2333-2340(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Hopkins R.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.M., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Rotherfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP SEQUENCE OF 11-188 FROM N.A.
RC TISSUE=Synovial sarcoma;
RX MEDLINE=95384157; PubMed=7655467;
RA de Leeuw B., Balleman M., Olde Weghuis D., Geurts van Kessel A.;
RT "Identification of two alternative fusion genes, SYT-SSX1 and
RT SYT-SSX2, in t(X;18)(p11.2;q11.2)-positive synovial sarcomas.";
RL Hum. Mol. Genet. 4:1097-1099(1995).
CC -!- TISSUE SPECIFICITY: Expressed at high level in the testis.
CC -!- FUNCTION: Could act as a modulator of transcription.
CC Expressed at low level in thyroid. Not detected in tonsil, colon,
CC lung, spleen, prostate, kidney, striated and smooth muscles.
CC Detected in rhabdomyosarcoma and fibrosarcoma cell lines. Not
CC detected in mesenchymal and epithelial cell lines.
CC -!- DISEASE: Participates in a t(X;18)(p11.2;q11.2) chromosomal
CC translocation that is specifically found in more than 80% of
CC synovial sarcoma and produces the SSX1-SSX1 or SSX1-SSX2 fusion
CC products. These hybrid proteins are probably responsible for
CC transforming activity. Heterogeneity in the position of the
CC breakpoint can occur (low frequency).
CC -!- SIMILARITY: Belongs to the SSX family.
CC -!- SIMILARITY: Contains 1 KRAB-related domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X86174; AAA01003.1; ..
EMBL; BC001003; AAA01003.1; ..

```
DR EMBL; S79325; AAB35378.1; -.
DR PIR; S55057; S55057.
DR Genew; HGNC:11335; SSX1.
DR MIM; 312820; -.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0003714; F:transcription co-repressor activity; TAS.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003655; KRAB_related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PS50806; KRAB RELATED; 1.
DR KX; Chromosomal translocation; Proto-oncogene, Multigene family;
KW Transcription regulation.
FT DOMAIN 20 83 KRAB-RELATED.
FT SITE 62 83 BREAKPOINT FOR TRANSLATION TO FORM THE
FT SITE 110 111 SSX1-SSX1 FUSION PROTEIN.
FT SITE 110 111 BREAKPOINT FOR TRANSLATION TO FORM THE
FT SITE 110 111 SSX1-SSX1 FUSION PROTEIN.
SQ SEQUENCE 188 AA; 21931 MW; E440D1B2AE3AE9F7 CRC64;

Alignment Scores:
Pred. No.: 8,96e-70 Length: 188
Score: 784.00 Matches: 149
Percent Similarity: 84.04% Conservative: 9
Best Local Similarity: 79.26% Mismatches: 30
Query Match: 76.49% Indels: 0
DB: 1 Gaps: 0

US-09-975-856-1 (1-576) x SSX1_HUMAN (1-188)
QY 1 ATGAAGCGAGACGACGCTTTGCAAGGAGACCCAGGATGATGCTCAAAATATCAGAGAAG 60
Db 1 MetAsnGlyAspThrPheAlaLysArgProArgAspAlaLysAlaSerGluLys 20
QY 61 TTACGAAAGCCTTCGATGATATGCAATATCTCTTAAGAAAGTGGGAAGATG 120
Db 21 ArgSerLysAlaPheAspAlaThrTyPheSerLysLysGluTrpLysMet 40
QY 121 AATCTCTCGAGAAATCGCTATGATATGATCAAGCTAACTATCATGAGTCTGACTAA 180
Db 41 LysTySerGluLysIleSerTyValTyMetLysArgAsnTyLysAlaMetThrLys 60
QY 181 CTAGTTTCAAGTCAACCTCCACCTTTCATCGGTAGTAAACGGCTCGACACTCCAC 240
Db 61 LeuGlyPheLysValThrLeuProProPheMetCysAsnLysGlnAlaThrAspPheGln 80
QY 241 GGAATGATTTTGTAACGATCGAAACACAGCAATCAGTTCACGCTCTCAGATGACT 300
Db 81 GlyAsnAspPheAspAsnAspHisAsnArgArgIleGlnValGluHisProGlnMetThr 100
QY 301 TTCGGCAGCTCCAGAGAATCTTCCCGAAGATCATGCCCAAGAACCCAGCAGGAAGAA 360
Db 101 PheGlyArgLeuHisArgIleProLysIleMetProLysLysProAlaGluAspGlu 120
QY 361 AATGTTTGAAGNAGTCCAGGATCTGCGCCCAAAATATGATGGAAACAGCTGTGC 420
Db 121 AsnAspSerLysGlyValSerGluAlaSerGlyProGlnAsnAspGlyLysGlnLeuHis 140
QY 421 CCCCCGGAAATCCAGTACCTTGGAGAAGATTAAACAGACATCTGACCCCAAGGGGG 480
Db 141 ProProGlyLysAlaAsnIleSerGluLysIleAsnLysArgSerGlyProLysArgGly 160
QY 481 AAACATCCTGGACCCACACACTCGTGGAGAAAGCAGCTGTGTTTATGAGAGATC 540
Db 161 LysHisAlaThrThrHisArgLeuArgGluArgLysGlnLeuValIleTyGluGluIle 180
QY 541 AGCCACCTCAGGAAGATGACGAG 564
Db 181 SerAspProGluGluAspAspGlu 188

RESULT 5
SSX5_HUMAN
ID SSX5_HUMAN STANDARD; PRT; 188 AA.
```

```
AC O60225; Q96AW3;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SSX5 protein.
GN SSX5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=98021352; PubMed=9378559;
RA Gure A.O., Tuercio O., Sahin U., Tsang S., Scanlan M.J., Jager E.,
RA Knuth A., Pfeundscht M., Old L.J., Chen Y.-T.;
RT "SSX: a multigene family with several members transcribed in normal
RT testis and human cancer.";
RL Int. J. Cancer 72:965-971(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udutin T.B., Toshiki S., Carninci P., Prange C.,
RA Rana S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting R., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Could act as a modulator of transcription.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O60225-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O60225-2; Sequence=VSP_006274;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Belongs to the SSX family.
CC -!- SIMILARITY: Contains 1 KRAB-related domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U0842; AAC05821.1; -.
DR EMBL; BC016640; AAH16640.1; -.
DR Genew; HGNC:11339; SSX5.
DR MIM; 300327; -.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003655; KRAB_related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PS50806; KRAB RELATED; 1.
KW Multigene family; Transcription regulation; Alternative splicing.
FT DOMAIN 20 83 KRAB-RELATED.
FT VARSPLIC 23 23 K -> KHPWRQVCDRGHVLNLSPPFKVGRFPASSIKALLC
```



```

Db 36 LysThrLeuTyrrAsnAspValMetGlnAspIleTyrrGluThrValIleSerLeuGlyLeu 55
QY 190 AAGTCCACCTCCACCTTTTCATCGTAGTAACAGGCTCCAGACTTCACGGGAATGAT 249
D 191 |||||
Db 56 LysLeuLys-----AsnAsp 60
QY 250 TTGGTACCATCGA----- 264
D 251 |||||
Db 61 ThrGlyAsnAspHisProIleSerValSerThrSerGluLeuGlnThrSerGlyCysGlu 80
QY 265 ---AACCACAGGATCAGGTTCAGTCGCTCAGATGACTTCGCG----- 306
D 266 |||||
Db 81 ValSerLysLysThrArgMetLysIleAlaGlnLysThrMetGlyArgGluAsnProGly 100
QY 307 -----AGCTCCAG-----AGATCTCCGAGATCATGCCAAGAGCCCA 348
D 308 |||||
Db 101 AspThrHisSerValGlnLysTrpHisArgAlaPheProArgLysLysArgLysLysPro 120
QY 349 GCAGAGGAAGAATGTTTTCAGGAAGTCCAGGATCTCGCCCAACAAATGATGGG 408
D 350 |||||
Db 121 AlaThrCysLysGlnGluLeuProLysLeuMetAspLeuHisGlyLysGlyProThrGly 140
QY 409 AACACAG 414
D 410 |||||
Db 141 GluLys 142

RESULT 7
Z398 HUMAN STANDARD; PRT; 642 AA.
AC Q8TD17; Q8TD18; Q9P2K7; Q9UDV8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 398 (Zinc finger DNA binding protein p52/p71).
GN ZNF398 OR ZER6 OR KIAA1339.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RP MEDLINE=21883946; PubMed=1179858;
RX Conroy A.T., Sharma M., Holtz A.E., Wu C., Sun Z., Weigel R.J.;
RA "A novel zinc finger transcription factor with two isoforms that are
RT differentially repressed by estrogen receptor-alpha.";
RL J. Biol. Chem. 277:9326-9334(2002).
RN [2]
SEQUENCE FROM N.A.
RP Bemis G., Langston Y., Tucci S.;
RA Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Sielen F.,
RA Dratchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Rahney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.F., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences."

```

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 234-642 FROM N.A.
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirokawa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:65-73(2000).
CC -!- FUNCTION: Function as a transcriptional activator.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=p71;
CC IsoId=Q8TD17-1; Sequence=Displayed;
CC Name=2; Synonyms=p52;
CC IsoId=Q8TD17-2; Sequence=VSP_006926;
CC -!- INDUCTION: By estrogen receptor alpha.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: Contains 1 KRAB domain.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.165-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AV049744; AAK92789.1; -
CC EMBL; AV049743; AAK92788.1; -
CC EMBL; AC004890; AAD45824.1; ALT_SEQ.
CC EMBL; EC043295; AAH43295.1; -
CC EMBL; AB037760; BAA92577.1; -
CC TRANSFAC; T05129; -
CC Genew; HGNC:18373; ZNF398.
CC GO; GO:0005634; C:nucleus; NAS.
CC GO; GO:0016563; P:transcriptional activator activity; NAS.
CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
CC InterPro; IPR001509; KRAB.
CC InterPro; IPR007087; Znf_C2H2.
CC Pfam; PF01352; KRAB; 1.
CC Pfam; PF00096; ZF-C2H2; 8.
CC ProDom; PD000003; Znf_C2H2; 2.
CC SMART; SM00349; KRAB; 1.
CC SMART; SM00355; Znf_C2H2; 8.
CC PROSITE; PS00805; KRAB; 1.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 7.
CC PROSITE; PS01157; ZINC_FINGER_C2H2_2; 8.
CC Transcription regulation; Activator; DNA-binding; Zinc-finger;
CC Metal-binding; Nuclear protein; Repeat; Alternative splicing.
CC -----
CC DOMAIN 143 214
CC KRAB.
CC C2H2-TYPE (ATYPICAL).
CC C2H2-TYPE (DEGENERATE).
CC C2H2-TYPE.
CC C2H2-TYPE.
CC C2H2-TYPE.
CC C2H2-TYPE.
CC C2H2-TYPE.
CC C2H2-TYPE.
CC Missing (in isoform 2).
CC /FTID=VSP_006926.
CC -----
CC SQ SEQUENCE 642 AA; 71311 MW; 69AA38FCD84FF633 CRC64;

```

```

Alignment Scores:
Pred. No.: 0.0593 Length: 642
Score: 101.00 Matches: 31
Percent Similarity: 41.60% Conservative: 21
Best Local Similarity: 24.80% Mismatches: 41

```



```
Db 139 SerLeuLeuMetGluAspIlePheGlyLysGluThrProSerGlyValThrMetGluArg 158
Qy 361 AATGGTTTGAAGGAA 375
Db 159 AlaGlyLeuGlyGlu 163

RESULT 9
PK3G_RAT
ID PK3G_RAT STANDARD; PRT; 1505 AA.
AC O70173;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phosphatidylinositol-4-phosphate 3-kinase C2 domain-containing gamma
DE polypeptide (EC 2.7.1.154) (Phosphoinositide 3-Kinase-C2-gamma)
DE (PtdIns-3-kinase C2 gamma) (PI3K-C2gamma).
GN PI3K2G.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
STAGE.
RC TISSUE=Regenerating liver;
RX MEDLINE=98184888; PubMed=9516481;
RA Ono F., Nakagawa T., Saito S., Owada Y., Sakagami H., Goto K.,
RA Suzuki M., Matsuno S., Kondo H.
RT "A novel class II phosphoinositide 3-kinase predominantly expressed in
RT the liver and its enhanced expression during liver regeneration."
RL J. Biol. Chem. 273:7731-7736(1998).
CC -!- FUNCTION: IN VITRO, PHOSPHORYLATES PTDINS AND PTDINS4P BUT NOT
CC PTDINS(4,5)P2.
CC -!- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol 4-
CC phosphate = ADP + 1-phosphatidyl-1D-myo-inositol 3,4-bisphosphate.
CC -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN NORMAL LIVER. HIGH
CC LEVELS ALSO FOUND IN REGENERATING LIVER. VERY LOW LEVELS FOUND IN
CC HEART AND TESTIS.
CC -!- DEVELOPMENTAL STAGE: HIGHER LEVELS OF EXPRESSION FOUND IN ADULT
CC LIVER THAN IN FETAL LIVER.
CC -!- SIMILARITY: Belongs to the PI3/P14-kinase family.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 1 pbox homology (PX) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB009636; BAA25634.1; .
CC InterPro; IPR008938; ARM.
CC InterPro; IPR000008; C2.
CC InterPro; IPR008973; C2_CaLB.
CC InterPro; IPR000403; PI3_P14_kinase.
CC InterPro; IPR002420; PI3K_C2.
CC InterPro; IPR000341; PI3K_ras_bind.
CC InterPro; IPR001263; PI3Ka.
CC InterPro; IPR001683; PX.
CC Pfam; PF00169; C2; 1.
CC Pfam; PF00454; PI3_P14_kinase; 1.
CC Pfam; PF00792; PI3K_C2; 1.
CC Pfam; PF00794; PI3K_rbd; 1.
CC Pfam; PF00613; PI3Ka; 1.
CC Pfam; PF00787; PX; 1.
CC SMART; SM00239; C2; 2.
CC SMART; SM00142; PI3K_C2; 1.
CC SMART; SM00144; PI3K_rbd; 1.
CC SMART; SM00145; PI3Ka; 1.
```

```
DR SMART; SM00146; PI3Kc; 1.
DR SMART; SM00312; PX; 1.
DR PROSITE; PS00499; C2 DOMAIN 1; FALSE NEG.
DR PROSITE; PS00004; C2 DOMAIN 2; FALSE NEG.
DR PROSITE; PS00915; PI3_4_KINASE_1; 1.
DR PROSITE; PS00916; PI3_4_KINASE_2; 1.
DR PROSITE; PS02900; PI3_4_KINASE_3; 1.
DR PROSITE; PS0195; PX; 1.
KW Transferase; Kinase; Membrane; Multigene family.
FT DOMAIN 976 1240 PI3K/PI4K.
FT DOMAIN 1259 1371 PX.
FT DOMAIN 1402 1499 C2 DOMAIN.
SQ SEQUENCE 1505 AA; 170974 MW; 5ED4C2329968C4B2 CRC64;

Alignment Scores:
Score: 0.95 Length: 1505
Pred. No.: 89.50 Matches: 33
Percent Similarity: 42.24% Conservative: 16
Best Local Similarity: 28.45% Mismatches: 50
Query Match: 8.89% Indels: 17
DB: 1 Gaps: 5

US-09-975-856-1 (1-576) x PK3G_RAT (1-1505)
Qy 490 AGGCATGTTTCCCTTTTGGTCCAGATGCTTGTAAATCTTCTCAAGTACTTGGAT 431
Db 83 ArgHisPheAsnGluPheThrSerGln-----SerProHisPheSerGln 97
Qy 430 TTCGGGGGGGACAGCTGTTTCCCATCATTTTGGGCCAGATGCTTGGCATTCCT 371
Db 98 LeuProPheGlyLysAlaSerAla-----IleGlyPheAsnProAlaValLeuPro 114
Qy 370 TCAACACCATTTCTCTCTCTGCTGCTTGGGCATGATCTTCGGGAAGATTCCTCGGA 311
Db 115 AlaHisGlnPheIleHisGluGlyAlaSerTrpArgAsnProThrArgLysTyrHisGly 134
Qy 310 GGCTGCCGAAGTCACTCTGAGGACGTTCAACCTGATTCCTGTGTTTTCGATCGT---TAC 254
Db 135 GlyGluAspProArgPheSerAlaLeuThrProSerSerThrGlyLeuAspLysCysHis 154
Qy 253 CAAATCATCTCCCGTGGAGTCTCAGCCGCTTACTACGCATGAAAGTGGGAGGTGA 194
Db 155 GlnGlnGlyGlnSerGlyThrGluHisCysAsnTyrTyrVal-----Glu 169
Qy 193 CTTGAAACCTAGTTTAGTACATGACCTCATGTTTAGCTTCATATACA 146
Db 170 ProGluAsnAsnVal-----ProHisHisTyrSerProTyrSer 182

RESULT 10
FMOL_CANFA
ID FMOL_CANFA STANDARD; PRT; 531 AA.
AC Q95LA2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dimethylalanine monooxygenase [N-oxide forming] 1 (EC 1.14.13.8)
DE (Hepatic flavin-containing monooxygenase 1) (FMO 1) (Dimethylalanine
DE oxidase 1).
DN FMOL.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Lattard V., Longin-Sauvageon C., Lachuer J., Burofossé T., Benoit E.;
RT "Cloning, sequencing and tissue dependent expression of FMO1 and FMO3
RT in the dog."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: This protein is involved in the oxidative metabolism of
CC a variety of xenobiotics such as drugs and pesticides.
CC -!- CATALYTIC ACTIVITY: N,N-dimethylaniline + NADPH + O(2) = N,N-
```

dimethylalanine N-oxide + NADP(+) + H(2)O.

COFACTOR: FAD (By similarity).

SUBCELLULAR LOCATION: Mitochondrion.

TISSUE SPECIFICITY: Liver.

SIMILARITY: Belongs to the FMO family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: AF384053; AAK97433.1; -

InterPro: IPR000759; Adrxn_reductase.

InterPro: IPR001327; FAD_Pyr_redox.

InterPro: IPR000960; Flav_cont_mnnoxgn.

PFam: PF00743; FMO-like; I.

PRINTS: PR00419; ADXRDPASE.

PRINTS: PR00368; FADPNR.

PRINTS: PR00370; FMOXYGENASE.

Oxidoreductase; Monooxygenase; NADP; Flavoprotein; FAD; Mitochondrion;

Transmembrane; Multigene family; Acetylation.

INIT MET 0 BY SIMILARITY.

MOD RES 1 1 ACETYLATION (BY SIMILARITY).

NP_BIND 8 13 FAD (ADP PART) (POTENTIAL).

NP_BIND 190 195 NADP (POTENTIAL).

SEQUENCE 531 AA; 59827 MW; 9F3458484540521F CRC64;

Alignment Scores:

Pred. No.:	0.996	Length:	531
Score:	88.50	Matches:	33
Percent Similarity:	38.21%	Conservative:	14
Best Local Similarity:	26.83%	Mismatches:	36
Query Match:	8.75%	Indels:	40
DB:	1	Gaps:	4

US-09-975-856-1 (1-576) x FMO1_CANFA (1-531)

QY 565 ACTCGTATCTCTCTCAGGGTCGCTGATCTCTTCATTAACACCCAGCGTCTTCTCTCAC 506

Db 43 ThrGluHisValGluGluGlyArgAlaSerLeuTyrIysSerVal----- 57

QY 505 GCAGTCTCTGGGTCGAGCGATGTTCCCGCTTTTGGGTCGAGATGTTCTTTAATCTCT 446

Db 58 -----ValSerAsnSerCysLysGluMetSerCys----- 67

QY 445 CCAAGGTACTTGGATTTCCCGGGGGCAGACACTGTTCCCATCATTTTGTGGCCAGATG 386

Db 68 -----TyrSerAspPro-----PheProGluAspTyr 77

QY 385 CCTCTGGCATTCTCTCAACACATTTCTCTCTGCTGCTGCTTCTGGGATGATCTCTCG 326

Db 78 ProAsnTyrValProAsnSerGluPheLeuGluTyrLeuLysMetTyrAla-Aen----- 95

QY 325 GGAAGATTCTCTGGAGGGTCCGGAAGTCATCTCAGGACGCTTCACCTGATCTCTGGT 266

Db 96 -----ArgPheSerLeuLeuLysCysII 103

QY 265 TTCGATCGTTACCAAAATCATCCCGTGGAGTCTGCGAGCCGCTTTACTACGATGAAG 206

Db 103 eArgPheLysThrLysValCysLysValThrLysCysProAspPheThrValThrGlyGI 123

QY 205 GTGGGAG 199

Db 123 ntrpGlu 125

RESULT 11

ID REPI_MOUSE

AC O54916; Q9C9J9; Q99LR8; PRT; 743 AA.

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DE 10-OCT-2003 (Rel. 42, Last annotation update)

DE RalBP1 associated Eps domain containing protein 1 (RalBP1-interacting protein 1).

DE REPS1.

GN Mus musculus (Mouse).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID:10090;

RN [1] SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.

RC TISSUE=Muscle.

RX MEDLINE=98058900; PubMed=9395447;

RA Yamaguchi A., Urano T., Goi T., Feig L.A.;

RT "An eps homology (EH) domain protein that binds to the ral-GTPase target, RalBP1."

RL J. Biol. Chem. 272:31230-31234 (1997).

RN [2]

RC SEQUENCE OF 8-743 FROM N.A. (ISOFORM 2).

RP STRAIN=C57BL/6J; TISSUE=Thymus;

RX MEDLINE=22354683; PubMed=12466851;

RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gofjabori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S., Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G., Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., King B.Z., Ringwald M., Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K., Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L., Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., Hirazane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y.;

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."

RL Nature 420:563-573 (2002).

RN [3]

RP SEQUENCE OF 447-743 FROM N.A. (ISOFORM 1).

RC TISSUE=Breast tumor;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler N.K., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ussid T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."

```
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
RN STRUCTURE BY NMR OF 227-318.
RX MEDLINE=21285759; PubMed=11389591;
RA Kim S., Cullis D.N., Reig L.A., Balleja J.D.;
RT "Solution structure of the Rep1 EH domain and characterization of
its binding to NP7 target sequences.";
RL Biochemistry 40:6776-6785(2001).
CC -!- FUNCTION: May coordinate the cellular actions of activated EGF
receptors and Ras-GTPases.
CC -!- SUBUNIT: Homodimer (Potential). Interacts with RALBP1, CRK and
GRB2. Binding to RALBP1 does not affect its Ras-binding activity.
CC Forms a complex with the SH3 domains of CRK and GRB2 which may
link it to an EGF-responsive tyrosine kinase.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O54916-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O54916-2; Sequence=VSP_007956; VSP_007957;
CC Note=Due to intron retention. No experimental confirmation
available;
CC -!- TISSUE SPECIFICITY: Expressed in all tissues examined. The highest
level expression was found in the kidney and testis.
CC -!- PTM: EGF stimulates phosphorylation on Tyr-residues.
CC -!- SIMILARITY: Contains 1 EH domain.
CC -!- SIMILARITY: Contains 1 EF-hand calcium-binding domain.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to a
frameshift in position 719.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
DR EMBL; AF031939; AAB94736.1; .
DR EMBL; AK041967; BAC31117.1; ALT INIT.
DR EMBL; BC002256; AAB02256.1; ALT_FRAME.
DR PIR; T09173; T09173.
DR PDB; 1FI6; 18-JUL-01.
DR MGD; MGI:1196373; Reps1.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; efhand; 1.
DR SMART; SM00027; EH; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00031; EH; 1.
KW Calcium-binding; Coiled coil; Phosphorylation; Alternative splicing;
3D-structure.
FT DOMAIN 233 324 EH.
FT CA_BIND 279 290 EF_HAND (POTENTIAL).
FT DOMAIN 488 551 PRO-RICH.
FT DOMAIN 599 743 INTERACTION WITH RALBP1.
FT DOMAIN 692 738 COILED COIL (POTENTIAL).
FT MOD_RES 236 236 PHOSPHORYLATION (POTENTIAL).
FT VARSPPLIC 368 402 QWTFSSRSSSTLTQFDSNIAPADPDTAIVHPV ->
VSKTSLSLISLFTGRSKQDRFTAGLYQYATP (in
isoform 2).
FT FTId=VSP_007956.
FT Missing (in isoform 2).
FT /FTId=VSP_007957.
SQ SEQUENCE 743 AA; 80598 MW; 25510D11254CF4A6 CRC64;
Alignment Scores:
Pred. No.: 1.18 Length: 743
Score: 88.00 Matches: 40
Percent Similarity: 38.92% Conservatives: 25
Best Local Similarity: 23.95% Mismatches: 41
Query Match: 8.59% Indels: 61
DB: 1 Gaps: 9
```

```
US-09-975-856-1 (1-576) x REP1_MOUSE (1-743)
QY 91 TACTTCTCTAAGAAAGAGTGGAAAGATGAATCTCGGAGAAAATCGTCTATGTGTAT 150
Db 11 TyrPheGlyArgSerGlnPhe-----Tyr 18
QY 151 ATGAAGCTAAACTATGAGTCAATGATCACTAACTAGTTTC----- 189
Db 19 IleAlaLeuLysLeuValAlaValAlaGlnSerGlyPheProLeuArgValGluSerIle 38
QY 190 -----AAGTCACCTCCACCTTCATCGCTAGTAAA----- 222
Db 39 AsnThrValLysAspLeuProleuProArgPheValAlaSerLysAsnGluGlnGluSer 58
QY 223 CGGGCTGCAGACTTCCACGGGAATGATTTTGGTAACGATCGAAACCCAC----- 270
Db 59 ArgLeuAlaAlaSerTyrSerSerAspSerGluAsnGlnGlySerTyrSerGlyValIle 78
QY 271 -----AGGAATCAGTTGAAAGTGAAGTCTCAGATGAGCTTCGGCAGCTCCAG 315
Db 79 ProProProProGlyArgGlyGlnValLysLysGlyProGlySerHisAspAlaValGln 98
QY 316 AGAATCTTCCCGAAGATCATGCCCAAGAGAGCCAGAGAGAGAAATGTTTGAAGAA 375
Db 99 -----ProArgProSerAlaGluGln----- 106
QY 376 GTGCCAGAGGCATCTGGCCCAAAATGATGGAAACAGCTGTGCCCCCGGGAATCCA 435
Db 107 -----GluProAlaSerProValValSerProGlnGln---SerProThrSerPro 123
QY 436 AGTACTCTGGAGAACATTAAACAAGACATCTGCACCAAAAGGGGAAACATG----- 487
Db 124 HisThr-TyrArgLysHisSerArgHis-----ProSerGlyLysSerGluArgPr 141
QY 488 -----CCTGCACCC 496
Db 141 OLeuThrGlyProGlyPro 147
RESULT 12
FMOL_PIG STANDARD; PRT; 531 AA.
ID AC FMOL_PIG
IC P16549;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dimethylalanine monooxygenase [N-oxide forming] 1 (EC 1.14.13.8)
DE (Hepatic flavin-containing monooxygenase 1) (FMO 1) (Dimethylalanine
oxidase 1).
DE FMOL OR FMO-1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]_TaxID=9823;
RP SEQUENCE FROM N.A., AND SEQUENCE OF 136-150 AND 308-317.
RC TISSUE=Liver;
RX MEDLINE=90212556; PubMed=2322534;
RA Gasser R., Tynes R.E., Lawton M.P., Korsmeyer K.K., Ziegler D.M.,
RA Philpot R.M.;
RT "The flavin-containing monooxygenase expressed in pig liver: primary
sequence, distribution, and evidence for a single gene.";
RL Biochemistry 29:119-124 (1990).
RN [2]
RP SEQUENCE OF 1-13 AND 184-201, AND ACETYLATION.
RC TISSUE=Liver;
RX MEDLINE=90343821; PubMed=2383273;
RA Guan S.H., Falick A.M., Cashman J.R.;
RT "N-terminus determination: FAD and NADP binding domain mapping of hog
liver flavin-containing monooxygenase by tandem mass spectrometry.";
RL Biochem. Biophys. Res. Commun. 170:937-943(1990).
RN [3]
RP SEQUENCE OF 185-207.
```

RC TISSUE=Liver;
RX MEDLINE=95278229; PubMed=7758472;
RA Wu R.-F. Ichikawa Y.;
RT An essential lysyl residue (Lys208) in the substrate-binding site of
RT porcine FAD-containing monooxygenase.;
RN Eur. J. Biochem. 229:749-753(1995).
RN [4]
RP CARBOHYDRATE-LINKAGE SITE.
RX MEDLINE=98451545; PubMed=9778310;
RA Korsmeyer K.K., Guan S., Yang Z.C., Falick A.M., Ziegler D.M.,
RA Cashman J.R.;
RT "N-glycosylation of pig flavin-containing monooxygenase form 1:
RT determination of the site of protein modification by mass
RT spectrometry.";
RL Chem. Res. Toxicol. 11:1145-1153(1998).
CC -!- FUNCTION: This protein is involved in the oxidative metabolism of
CC a variety of xenobiotics such as drugs and pesticides.
CC -!- CATALYTIC ACTIVITY: N,N-dimethylaniline + NADPH + O(2) = N,N-
CC dimethylaniline N-oxide + NADP(+) + H(2)O.
CC -!- COFACTOR: FAD.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- TISSUE SPECIFICITY: Liver.
CC -!- SIMILARITY: Belongs to the FMO family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M32031; AAA31033.1; -;
CC PIR: A33768; A33768.
DR InterPro: IPR000759; Adrnx_reductase.
DR InterPro: IPR001327; FAD_pyr_redox.
DR InterPro: IPR000960; Flav_cont_mnoxgn.
DR Pfam: PF00743; FMO-like; 1.
DR PRINTS: PR00419; ADXEDTASE.
DR PRINTS: PR00368; FADPNR.
DR PRINTS: PR00370; FMOXYGENASE.
KW Oxidoreductase; Monooxygenase.
KW Transmembrane; Multigene family; Acetylation; Glycoprotein.
FT INIT MET 0
FT MOD_RES 1 0
FT NP_BIND 8 13 ACETYLATION.
FT NP_BIND 190 195 FAD (ADP PART) (POTENTIAL).
FT ACT_SITE 207 207 NADP (POTENTIAL).
FT ACT_SITE 207 207 SUBSTRATE BINDING.
FT CARBOHYD 119 119 N-LINKED GLCNAC... (HIGH MANNOSIDE).
SQ SEQUENCE 531 AA; 59821 MW; 58475E6D81C8157 CRC64;

Alignment Scores:
Pred. No.: 1.25 Length: 531
Score: 87.50 Matches: 33
Percent Similarity: 37.40% Conservative: 13
Best Local Similarity: 26.83% Mismatches: 37
Query Match: 9.69% Indels: 40
DB: 1 Gaps: 4

US-09-975-856-1 (1-576) x FMO1_PIG (1-531)

QY 565 ACTCGTCTCTCTCTCAGGTCGTCGATCTCTTCATTAACACACAGCTGCTTCTCTCAC 506
Db 43 ThrGluHisValGluGluGlyArgAlaSerLeuTyrLysSerVal----- 57
QY 505 GCAGCTCTGGTCCAGGCATGTTCCCCCTTTTGGGTCCAGAGTCTTGTGTAATCTTCT 446
Db 58 -----ValSerAsnSerCysLysGluMetSerCys-----° 67
QY 445 CCAAGGTACTGGATTCTCCGGGGGACAGCTGTTTCCCATCATTTTGTGGGCGCAGATG 386
Db 69 -----TyrProAspPhePro-----PheProGluAspTyr 77

QY 385 CCTCTGGCACTTCTCTCAACACATTTCTCTCTGCTGGTCTTCTTGGGCATGATCTTCG 326
Db 78 ProAsnTyrValProAsnSerHisPheLeuGluTyrLeu----- 90
QY 335 GGAAGATTCTCTGGAGGCTGCCGAAAGTCATCTGAGGACGTTCAACCTGATCTCTGGT 266
Db 91 -----Arg-MetTyrAlaAsnGlnPheAsnLeuLeuLysCysI 103
QY 265 TTCGATCGTTACCAAAATCATTCCGTCGGAAGTCTGCAGCCGCTTTACTAGCATGAAG 206
Db 103 eGlnPheLysThrLysValCysSerValThrLysHisGluAspPheAsnThrThrGlyG 123
QY 205 GTGGGAG 199
Db 123 nTrpAsp 125

RESULT 13
REPL_HUMAN STANDARD; PRT; 744 AA.
AC Q96D71; Q8NDR7; Q8WU62; Q9BX99;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE RalBP1 associated Eps domain containing protein 1 (RalBP1-interacting
DE protein 1).
GN REPS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Mao Y., Xie Y., Zhou Z., Zhao W., Zhao S., Wang W., Huang Y., Wang S.,
RA Tang R., Chen X., Wu C.;
RL Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).
RC TISSUE=Brain, and Lymph;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Prange C.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 274-743 FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RA Bioecker H., Boecker M., Brandt P., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (Jul-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May coordinate the cellular actions of activated EGF
CC receptors and Ral-GTPases (By similarity).
CC -!- SUBUNIT: Homodimer (potential). Interacts with RALBP1, CRK and
CC GRB2. Binding to RALBP1 does not affect its Ral-binding activity.
CC Forms a complex with the SH3 domains of CRK and GRB2 which may
CC link it to an EGF-responsive tyrosine kinase (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=3;
CC Name=1;

```

CC      IsoId=Q96D71-1; Sequence=Displayed;
CC      Names=2;
CC      IsoId=Q96D71-2; Sequence=VSP_007953, VSP_007954;
CC      Names=3;
CC      IsoId=Q96D71-3; Sequence=VSP_007955;
CC      -|- PM: EGF stimulates phosphorylation on Tyr-residues (By
CC      similarity);
CC      -|- SIMILARITY: Contains 1 EH domain.
CC      -|- SIMILARITY: Contains 1 EF-hand calcium-binding domain.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF251052; AAK34942.1; -
DR      EMBL; BC012764; AAH12764.1; -
DR      EMBL; BC021211; AAH21211.1; -
DR      EMBL; AL331900; CAD38569.1; -
DR      Genew; HGNC:15578; REPS1.
DR      InterPro; IPR002048; EF-hand.
DR      InterPro; IPR002048; EFS15_homology.
DR      SMART; SM00027; EH_1.
DR      SMART; SM00018; EF_HAND; 1.
DR      PROSITE; PS00033; EH_1.
DR      PROSITE; PS00033; EH_1.
KW      Calcium-binding; Coiled coil; Phosphorylation; Alternative splicing.
FT      DOMAIN 233 324
FT      CA_BIND 279 290
FT      DOMAIN 489 552
FT      DOMAIN 600 744
FT      DOMAIN 693 739
FT      MOD_RES 236 236
FT      VARSPIC 368 394
FT      VARSPIC 458 521
FT      VARSPIC 458 521
FT      VARSPIC 458 521
FT      VARSPIC 458 521
FT      CONFLICT 576 576
FT      CONFLICT 665 665
FT      CONFLICT 742 743
SQ      SEQUENCE 744 AA; 80769 MW; 1DFF29711DB2B54 CRC64;

Alignment Scores:
Pred. No.: 1.49 Length: 744
Score: 87.00 Matches: 39
Percent Similarity: 38.89% Conservatives: 24
Best Local Similarity: 24.07% Mismatches: 48
Query Match: 8.49% Indels: 51
DB: 1 Gaps: 9

US-09-975-856-1 (1-576) x REP1_HUMAN (1-744)
Qy 91 TACTTCTTAAGAAGAGTGGGAAAGATGAATCTCGGAGAAAATCGTCTATGCTAT 150
Db 11 TyrPheGlyArgSerGlnPhe-
Qy 151 ATGAAGCTAACTATGAGTCTAGCTAACTAGGTTTC----- 189
Db 19 IleAlaLeuLysLeuValAlaValAlaGlnSerGlyPheProLeuArgValGluSerIle 38
Qy 190 -----AAGGTCCCTCCACCTTTTCAGTGTAGTAA----- 222
Db 39 AsnThrValLysAspLeuProLeuProArgPheValAlaSerLysAsnGluGlnGluSer 58
Qy 223 CGGGCTGCAGACTTCCACGGGAATCATTTTGGTAACGATCGAACCACAGGAATCAGGTT 282
Db 59 ArgHisAlaAlaSerTyrSerAspSerGluAsnGlnGlnGlySerTyrSerGlyValIle 78

```

```

Qy 283 GAACGTCTCT-----CAGATGACTTTCCGGCGCTCCAGAGAAATCTTCCCGAAG 330
Db 79 PropProProGlyArgGlyGlnValLysLysGlySerValSerHis-----AspThr 96
Qy 331 ATCATGCCCAAGAACCCAGACAGAGAGAAAATGGTTTGAAGGAAGTCCAGAGGCATCT 390
Db 97 ValGlnProArgThrSerAlaAspAlaGln-----GluProAla 109
Qy 391 GSCCCCAAAAATGATGGGAAACAGCTGTGCCCCCGGGAATCCAAAGTACCTTGAGAG 450
Db 110 SerProValValSerProGlnGln---SerProProThrSerProHisThr-TripArgLy 128
Qy 451 ATTAACAAGACATCTGGACCCCAAAAGGGGAAACATG-----CCTGG 492
Db 128 HisSerArgHis-----ProSerGlyGlyAsnSerGluArgProLeuAlaGlyProGl 146
Qy 493 ACCC 496
Db 146 yPro 147

RESULT 14
SK11_CHICK
ID SK11_CHICK STANDARD; PRT; 396 AA.
AC P48435;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Transcription factor SOX-11.
GN SOX11.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=95267693; PubMed=7748786;
RA Uwanogho D., Rex M., Cartwright E.J., Pearl G., Healy C.,
RA Scotting P.J., Sharpe P.T.;
RT "Embryonic expression of the chicken Sox2, Sox3 and Sox11 genes
RT suggests an interactive role in neuronal development.";
RL Mech. Dev. 49:23-36(1995).
CC -|- FUNCTION: May function as switches in neuronal development.
CC -|- SUBCELLULAR LOCATION: Nuclear.
CC -|- TISSUE SPECIFICITY: Low level expression is seen in
CC undifferentiated proliferating cells of neural epithelium. A
CC greater expression is seen in the maturing neurons after they
CC leave the neural epithelium. It is also found in the gut
CC epithelium and adrenal medulla.
CC -|- SIMILARITY: Contains 1 HMG box domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR      EMBL; U12534; AAB09664.1; -
DR      PIR; I50707; I50707.
DR      HSSP; P48436; LSX9.
DR      InterPro; IPR000910; HMG_12_box.
DR      Pfam; PF00505; HMG_box; 1.
DR      SMART; SM00398; HMG; 1.
DR      PROSITE; PS0118; HMG_BOX_2; 1.
KW      Transcription regulation; DNA-binding; Nuclear protein.
FT      DNA_BIND 49 117 HMG_BOX.
FT      DOMAIN 162 155 POLY-ALA.
FT      DOMAIN 284 212 POLY-GLU.
FT      DOMAIN 288 294 POLY-PRO.
FT      DOMAIN 332 335 POLY-ARG.

```



```
DR EMBL; X07290; CAA30269.1; -.
DR PIR; S00754; S00754.
DR Genew; HGNC:13104; ZNF38.
DR Genew; HGNC:13089; ZNF3.
DR MIM; 601261; -.
DR MIM; 194510; -.
DR GO; GO:0005634; Cnucleus; IC.
DR GO; GO:0003700; F:transcription factor activity; NAS.
DR GO; GO:0008270; F:zinc ion binding; NAS.
DR GO; GO:0045321; P:cell activation; NAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR007086; Znf_C2H2_sub.
DR Pfam; PF01352; KRAB; 1.
DR Pfam; PF00096; zf-C2H2; 8.
DR PRINTS; PR00048; ZINCFINGER.
DR ProDom; PD000003; Znf_C2H2; 8.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 8.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 8.
DR PROSITE; PS0805; KRAB; 1.
DR PROSITE; PS0805; KRAB; 1.
KW Transcription regulation; Activator; Zinc-finger; DNA-binding;
KW Repeat; Metal-binding; Nuclear protein; Differentiation.
FT DOMAIN 51 123
FT ZN_FING 200 227
FT ZN_FING 228 255
FT ZN_FING 256 283
FT ZN_FING 284 311
FT ZN_FING 312 339
FT ZN_FING 340 367
FT ZN_FING 368 395
FT ZN_FING 396 423
FT ZN_FING 421 21
FT CONFLICT 252 256
FT CONFLICT 336 387
FT CONFLICT 336 387
SQ SEQUENCE 446 AA; 50932 MW; 67A5926807304782 CRC64;

Alignment Scores:
Pred. No.: 1.53 Length: 446
Score: 86.50 Matches: 42
Percent Similarity: 36.98% Conservative: 29
Best Local Similarity: 21.88% Mismatches: 50
Query Match: 8.44% Indels: 71
DB: 1 Gaps: 8

US-09-975-856-1 (1-576) x ZN38_HUMAN (1-446)
QY 40 GATGCTCAATATACAGAGAGTTACGAAAGGCC-----TTGGAT 78
Db 35 AspGluMetLeuAlaAlaLeuLeuLysSerGlnGluValThrPheGlu 54
QY 79 GATATTGCCAATACTCTCTAAGAGAGTGGGAAAAGATGAATCCTCGGAGAAAATC 138
Db 55 AspValAlaValThrPheLeuArgLysGluTrpLysArgLeuGluProAlaGlnArgAsp 74
QY 139 GTCTATGTGTATGAAGCTA---AACTATGAGGTCATGACTAAACTAGGTTTCAAGGTC 195
Db 75 LeuTyArgAspValMetLeuGluAsnTyr----- 84
QY 196 ACCCTCCACCTTTCATGCGTAGCTAAACGGGCTGCAGACTTCACGGGAATGATTTGGT 255
Db 85 -----GlyAsnValPheSer 89
QY 256 AACGATCGAACACACAGG-----AATCAGGTTGAACGTCCTCAGATGACTTTC 303
Db 90 LeuAspArgGluThrArgThrGluAsnAspGlnGluIleSerGluAspThrArgSerHis 109
QY 304 GGCGGCTCCAGAGATCTTCCGAAG-----ATCATGCCAAGAGGCCA 348
Db 304 -----ATCATGCCAAGAGGCCA 348
```

```
Db 110 GlyValLeuLeuGlyArgPheGlnLysAspIleSerGlnGlyLeuLysPheLysGluAla 129
QY 349 GCAGAGGAAGAAAATGTTTCAAGAGAGTCCAGAGGCATCTGCCCCACAAAATGATGGG 408
Db 130 TyrGluArgGluValSerLeuLysArg----- 138
QY 409 AACAGCTGTGCCCCCGGGAATCCAAATCCAACTTGGAGAGAGATTAAACAAGACATCTGGA 468
Db 139 -----ProLeuGlyAsn---SerProGlyGluArgLeuAsnArg----- 150
QY 469 CCCAAAAGGGGAAACATGCTGTGGACCCACAGACTGCGTGAGAGAAAAGCAGCTGGTGT 528
Db 151 -----LysMetProAspPheGlyGlnValThrVal 160
QY 529 TATGAAGAGATCAGCGACCCCTGAGGAAGATGACGAG 564
Db 161 GluGluLysLeuThrProArgGlyGluArgSerGlu 172

Search completed: March 31, 2004, 13:58:04
Job time : 18 secs
```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 31, 2004, 13:40:37 ; Search time 47 Seconds
(without alignments)
7733.553 Million cell updates/sec

Title: US-09-975-856-1

Perfect score: 1025

Sequence: 1 ATGACGAGGACGACGCCTT.....ATGACGAGTAACCTCCCTCG 576

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-Q/csn2_1/USFTO_spool_3/US975856/runat_31032004_132903_14748/app_query.fasta_1.775
-DB=SPREMBL_25 -QFMT=fastaan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US975856 @CGN 1.1.86 @runat_31032004_132903_14748 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG-SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPREMBL_25:
1: sp_archaea*
2: sp_bacteria*
3: sp_fungi*
4: sp_human*
5: sp_invertebrate*
6: sp_mammal*
7: sp_mhc*
8: sp_organelle*
9: sp_phage*
10: sp_plant*
11: sp_rodent*
12: sp_virus*
13: sp_vertebrate*
14: sp_unclassified*
15: sp_rvirus*
16: sp_bacteriap*
17: sp_archaeap*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	685.5	66.9	223	4	Q9BU88 homo sapien

ID	Q9BU88	PRELIMINARY;	PRT;	223 AA.
AC	Q9BU88;			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	Similar to synovial sarcoma, X breakpoint 2.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	Tissue=Placenta;			
RA	Strausberg R.;			
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC002818; AA02818.1; -			
DR	GO; GO:0005622; C:intracellular; IEA.			
DR	GO; GO:0003676; F:nucleic acid binding; IEA.			
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.			
DR	InterPro; IPR001909; KRAB.			
DR	InterPro; IPR003655; KRAB_related.			
DR	Pfam; PF01352; KRAB; 1.			
DR	SMART; SM00349; KRAB; 1.			

	2	636	62.0	170	4	Q9BRW7	Q9brw7 homo sapien
3	576	56.2	155	4	Q96011	Q96q11 homo sapien	
4	545.5	53.2	136	4	Q96010	Q96q10 homo sapien	
5	363	35.4	98	4	Q9Y444	Q9y444 homo sapien	
6	255	24.9	64	4	Q8WZ29	Q8wz29 homo sapien	
7	247.5	24.1	113	4	Q9NZK4	Q9nzK4 homo sapien	
8	214.5	20.9	128	11	Q80ZT4	Q80zT4 mus musculus	
9	214.5	20.9	165	11	Q8C5Z3	Q8c5Z3 mus musculus	
10	211.5	20.6	170	11	Q9CPU1	Q9cpU1 mus musculus	
11	160.5	15.7	117	4	O75101	O75101 homo sapien	
12	137	13.4	39	4	O81ZHO	O81zh0 homo sapien	
13	137	13.4	39	4	O81ZG9	O81zg9 homo sapien	
14	137	13.4	39	4	O81ZG8	O81zg8 homo sapien	
15	137	13.4	39	4	O81ZG7	O81zg7 homo sapien	
16	137	13.4	39	4	O81ZG6	O81zg6 homo sapien	
17	137	13.4	39	4	O81ZG5	O81zg5 homo sapien	
18	124	12.1	39	4	O81ZH1	O81zh1 homo sapien	
19	113.5	11.1	510	4	O86TD5	O86td5 homo sapien	
20	112.5	11.0	1111	4	O60290	O60290 homo sapien	
21	111	10.8	330	4	Q96XN0	Q96xn0 homo sapien	
22	110.5	10.8	524	4	Q8N2J5	Q8n2J5 homo sapien	
23	110.5	10.8	785	4	Q9ULD5	Q9uld5 homo sapien	
24	109.5	10.7	652	11	Q8BFS8	Q8bfS8 mus musculus	
25	109	10.6	432	3	Q9HF87	Q9hf87 botrytis ci	
26	109	10.6	635	11	Q8QZ22	Q8qz22 mus musculus	
27	106.5	10.4	782	4	Q8N393	Q8n393 homo sapien	
28	106	10.3	579	11	Q9NVI0	Q9nvi0 mus musculus	
29	104.5	10.2	405	4	Q8IW91	Q8iW91 homo sapien	
30	104.5	10.2	579	11	Q8C393	Q8c393 mus musculus	
31	104	10.1	643	11	Q8BV16	Q8bv16 mus musculus	
32	103	10.0	614	11	Q8C887	Q8c887 mus musculus	
33	103	10.0	614	11	Q8BPJ3	Q8bpj3 mus musculus	
34	103	10.0	621	11	Q8OTC5	Q8otC5 mus musculus	
35	102	10.0	511	11	Q8BVH0	Q8bvH0 mus musculus	
36	102	10.0	649	4	Q9P215	Q9p215 homo sapien	
37	100.5	9.8	182	4	Q9UDV5	Q9udv5 homo sapien	
38	100.5	9.8	431	11	Q8C964	Q8c964 mus musculus	
39	98.5	9.6	258	17	O8TPS6	O8tps6 methanosarc	
40	98.5	9.6	428	17	O8TUH2	O8tuh2 methanosarc	
41	98.5	9.6	428	17	O8TYT7	O8tyT7 methanosarc	
42	98.5	9.6	428	17	O8TNA0	O8tna0 methanosarc	
43	98.5	9.6	428	17	O8TMC5	O8tmc5 methanosarc	
44	98.5	9.6	428	17	O8TL32	O8tl32 methanosarc	
45	98.5	9.6	428	17	Q8TIL4	Q8til4 methanosarc	

ALIGNMENTS

RESULT 1

ID	Q9BU88	PRELIMINARY;	PRT;	223 AA.
AC	Q9BU88;			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	Similar to synovial sarcoma, X breakpoint 2.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	Tissue=Placenta;			
RA	Strausberg R.;			
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC002818; AA02818.1; -			
DR	GO; GO:0005622; C:intracellular; IEA.			
DR	GO; GO:0003676; F:nucleic acid binding; IEA.			
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.			
DR	InterPro; IPR001909; KRAB.			
DR	InterPro; IPR003655; KRAB_related.			
DR	Pfam; PF01352; KRAB; 1.			
DR	SMART; SM00349; KRAB; 1.			

DR PROSITE; PS50806; KRAB RELATED; 1.
SQ SEQUENCE 223 AA; 25173 MW; 2BF8B1FFA4D58094 CRC64;
Alignment Scores:
Pred. No.: 1.15e-62 Length: 223
Score: 685.50 Matches: 137
Percent Similarity: 66.07% Conservative: 11
Best Local Similarity: 61.16% Mismatches: 26
Query Match: 66.88% Indels: 50
DB: 4 Gaps: 1

US-09-975-856-1 (1-576) x Q9BU88 (1-223)

QY 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATCTCAATATCAGAGAA 60
DB 1 MetAsnGlyAspAlaPheAlaArgProThrValGlyAlaGlnIleProGluLys 20
QY 61 TTACGAAAGCCCTTCGATGATATTGCCAAATCTCTTAAGAAAGAGTGGGAAGAGTG 120
DB 21 IleglnLysAlaPheAspAlaLysPheSerLysGluGlnTrpGluLysMet 40
QY 121 AATCTCGGAGAAATCGCTATGTATATGAAGCTAACTATCAGGTCATGACTAAA 180
DB 41 LysAlaSerGluLysIlePheTyrValTyrMetLysArgLysTyrGluAlaMetThrLys 60
QY 181 CTAGGTTTCAAGTCACCTCCACCTTCATCGCTAGTAAACGGGCTCAGACTCCAC 240
DB 61 LeuGlyPheLysAlaThrLeuProPheMetCysAsnLysArgAlaGluAspPheGln 80
QY 241 GGAATGATTTCGTACGATCGAAACACACGAGGATCAGGTTCAACGCTCTCAGATGACT 300
DB 81 GlyAsnAspLeuAspAsnAspProAsnArgGlyAsnGlnValGluArgProGlnMetThr 100
QY 301 TTCCGACGCTCCAGAGAACTCTCCGAGATCATGCCAGAGCCAGACAGAGAGAA 360
DB 101 PheGlyArgLeuGlnValProGluAlaSerGlyProGlnAsnAspGlyLysGluLeuCys 120
QY 361 AATGCTTTGAAGAAAGTCCAGAGGATCTGGCCCAAAATCATGGGAAACAGCTGTCC 420
DB 421 CCCCCGGAATCCAGTACTCTGGAGAGATTAAACAGACATCTCGA 468
QY 141 ProGlyLysProThrThrSerGluLysIleHisGluArgSerGlyAsnArgGluAla 160
QY 468 ----- 468
DB 161 GlnGluLysGluGluArgArgGlyThrAlaHisArgTrpSerSerGlnAsnThrHisAsn 180
QY 468 ----- 468
DB 181 IleGlyArgPheSerLeuSerThrSerMetGlyAlaValHisGlyThrProLysThrIle 200
QY 469 -----CCCAAAGGGGAACATGCTGGACCCACAGACTGGTGAGAGA 513
DB 201 ThrHisAsnArgAspPro-LysGlyGlyAsnMetProGlyProThrAspCysValArgG 220
QY 514 AAGCAGCTGG 523
DB 220 uAsnSerTrp 223

RESULT 2

Q9BRW7 PRELIMINARY; PRT; 170 AA.

ID Q9BRW7; (TrEMBLrel. 17, Created)
AC Q9BRW7; (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 25, Last annotation update)
DE Synovial sarcoma, X breakpoint 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.
RC TISSUE-Bone marrow;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005904; AA05904.1; -
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003676; P:nucleic acid binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003655; KRAB-related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PS50806; KRAB RELATED; 1.
SQ SEQUENCE 170 AA; 19457 MW; D981F807A9C7EAB1 CRC64;

Alignment Scores:

Pred. No.: 1.61e-57 Length: 170
Score: 636.00 Matches: 122
Percent Similarity: 82.05% Conservative: 6
Best Local Similarity: 78.21% Mismatches: 28
Query Match: 62.05% Indels: 0
DB: 4 Gaps: 0

US-09-975-856-1 (1-576) x Q9BRW7 (1-170)

QY 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATCTCAATATCAGAGAA 60
DB 1 MetAsnGlyAspAlaPheAlaArgProThrValGlyAlaGlnIleProGluLys 20
QY 61 TTACGAAAGCCCTTCGATGATATTGCCAAATCTCTTAAGAAAGAGTGGGAAGAGTG 120
DB 21 IleglnLysAlaPheAspAlaLysPheSerLysGluGlnTrpGluLysMet 40
QY 121 AATCTCGGAGAAATCGCTATGTATATGAAGCTAACTATGAGTCAAGTCACTGACTAAA 180
DB 41 LysValSerGluLysIleValTyrValTyrMetLysArgLysTyrGluAlaMetThrLys 60
QY 181 CTAGGTTTCAAGTCACCTCCACCTTCATCGCTAGTAAACGGGCTCAGACTCCAC 240
DB 61 LeuGlyPheLysAlaThrLeuProPheMetCysAsnLysArgAlaGluAspPheGln 80
QY 241 GGAATGATTTCGTACGATCGAAACACACGAGGATCAGGTTCAACGCTCTCAGATGACT 300
DB 81 GlyAsnAspLeuAspAsnAspProAsnArgGlyAsnGlnValGlnArgProGlnMetThr 100
QY 301 TTCCGACGCTCCAGAGAACTCTCCGAGATCATGCCAGAGCCAGACAGAGAGAA 360
DB 101 PheGlyArgLeuGlnGlyIlePheProLysIleMetProLysLysProAlaGluGluGly 120
QY 361 AATGCTTTGAAGAAAGTCCAGAGGATCTGGCCCAAAATCATGGGAAACAGCTGTCC 420
DB 121 AsnValSerLysGluValProGluAlaSerGlyProGlnAsnAspGlyLysGlnLeuCys 140
QY 421 CCCCCGGAATCCAGTACTCTGGAGAGATTAAACAGACATCTGGA 468
DB 141 ProGlyLysProThrThrSerGluLysIleAsnMetIleSerGly 156

RESULT 3

Q9EQ11 PRELIMINARY; PRT; 155 AA.

ID Q9EQ11; (TrEMBLrel. 19, Created)
AC Q9EQ11; (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 25, Last annotation update)
DE DJ54B20.1.1 (Novel SSX family protein (isoform 1)) (Fragment).
GN DJ54B20.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.

RA Grafham D.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z98304; CAC41946.1; -.
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001909; KRAB.
DR InterPro: IPR003655; KRAB_related.
DR Pfam: PF01352; KRAB; 1.
DR SMART: SM00349; KRAB; 1.
DR PROSITE: PS0806; KRAB_RELATED; 1.
FT NON_TER 155
SQ SEQUENCE 155 AA; 17689 MW; 06EC7C9D935A1A30 CRC64;
Alignment Scores:
Pred. No.: 2,9e-51 Length: 155
Score: 576.00 Matches: 113
Percent Similarity: 80.65% Conservative: 12
Best Local Similarity: 72.90% Mismatches: 30
Query Match: 56.20% Indels: 0
DB: 4 Gaps: 0
US-09-975-856-1 (1-576) x Q96Q10 (1-155)
QY 1 ATGAACGGAGACGCGCTTTGCAAGGAGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
Db 1 MetAsnGlyAspAlaPheAlaLysArgProArgAspAlaLysArgLys 20
QY 61 TTACGAAAGCCTTCGATATATGCAAAATCTTCTTAAGAAAGATGGGAAAGATG 120
Db 21 ArgSerLysAlaPheAspAlaLysArgProArgAspAlaLysArgLysMet 40
QY 121 AAATCTTCGAGAAATCGCTATGCTATATCAAGCTAACTATGAGTCACTATAA 180
Db 41 LysPheSerGluLysIleSerCysValHisMetLysArgLysTyrGluAlaMetThrLys 60
QY 181 CTAGGTTTCAAGGTCACCTCCACCTTCATGCTAGTAAAGGCTGAGACTTCCAC 240
Db 61 LeuGlyPheAsnValThrLeuSerLeuPheMetArgAsnLysArgAlaThrAspSerGln 80
QY 241 GGGAAATGTTTGTACGATCGAAGACACAGGATCAGTGTAAAGTCTCAGATGACT 300
Db 81 ArgAsnAspSerAspAsnAspArgAsnArgGlyAsnGluValGluArgProGlnMetThr 100
QY 301 TTCCGACGCTCCAGAGATCTTCCCGAAGATCATGCCAAGAGCCAGCAGAGGAGAA 360
Db 101 PheGlyArgLeuGlnArgIlePheProLysIleMetProGluLysProAlaGluGly 120
QY 361 AATGGTTTGAAGGAGTCCAGAGGATCTGGCCCAAAATATGATGGGAAACAGCTGTC 420
Db 121 SerAspSerLysGlyValProGluAlaSerGlyProGlnAsnAspGlyLysLysLeuCys 140
QY 421 CCCCCGGAAATCCAGTACCTTGGAGAAGATTAAACAGACATCT 465
Db 141 ProProGlyLysAlaSerSerSerGluLysIleHisGluArgSer 155
RESULT 4
Q96Q10 PRELIMINARY; PRT; 196 AA.
AC Q96Q10;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DJS4B20.1.2 (Novel SSX family protein (isoform 2)) (Fragment).
GN DJS4B20.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Grafham D.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: Z98304; CAC41947.1; -.
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0003676; F:nucleic acid binding; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001909; KRAB.
DR InterPro: IPR003655; KRAB_related.
DR Pfam: PF01352; KRAB; 1.
DR SMART: SM00349; KRAB; 1.
DR PROSITE: PS0806; KRAB_RELATED; 1.
FT NON_TER 196
SQ SEQUENCE 196 AA; 22299 MW; 71CALBE1F4BA07F8 CRC64;
Alignment Scores:
Pred. No.: 4,59e-48 Length: 196
Score: 545.50 Matches: 113
Percent Similarity: 63.78% Conservative: 12
Best Local Similarity: 57.65% Mismatches: 30
Query Match: 53.22% Indels: 41
DB: 4 Gaps: 1
US-09-975-856-1 (1-576) x Q96Q10 (1-196)
QY 1 ATGAACGGAGACGCGCTTTGCAAGGAGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
Db 1 MetAsnGlyAspAlaPheAlaLysArgProArgAspAlaLysArgLys 20
QY 61 TTACGAAAGCCTTCGATATATGCAAAATCTTCTTAAGAAAGATGGGAAAGATG 120
Db 21 ArgSerLysHisProTyrArgLysValCysAspLeuAlaLeuHisLeuValThrLeuThr 40
QY 69 ----- 69
Db 41 ProPheTyrLysValGlyArgGluProAlaSerIleThrGluAlaLeuLeuCysGlyArg 60
QY 70 ----- GCCTTCGATGATATGCAAAATCTTCTTAAGAAAGATGGGAAAGATG 117
Db 61 GlyGluAlaArgAlaPheAspAlaLysArgProArgAspAlaLysArgLysGluTyrGluLys 80
QY 118 ATGAATCTTCGAGAAATCGCTATGCTATATGAACTAACTATGAGTCACTACT 177
Db 81 MetLysPheSerGluLysIleSerCysValHisMetLysArgLysTyrGluAlaMetThr 100
QY 178 AAATGAGTTTCAAGTCCACCTCCACCTTCATGCTAGTAAAGGCTGAGACTTCC 237
Db 101 LysLeuGlyPheAsnValThrLeuSerLeuPheMetArgAsnLysArgAlaThrAspSer 120
QY 238 CACGGAGATGATTTGGTAACGATCGAAGACACAGGATCAGGTTGACCTCCAGATG 297
Db 121 GlnArgAsnAspSerAspAsnAspArgAsnArgGlyAsnGluValGluArgProGlnMet 140
QY 298 ACTTTCCGACGCTCCAGAGATCTTCCCGAAGATCATGCCAAGAGCCAGCAGAGGAA 357
Db 141 ThrPheGlyArgLeuGlnArgIlePheProLysIleMetProGluLysProAlaGluLys 160
QY 358 GAAATGTTTGAAGGAGTCCAGAGGATCTGGCCCAAAATATGATGGGAAACAGCTG 417
Db 161 GlySerAspSerLysGlyValProGluAlaSerGlyProGlnAsnAspGlyLysLysLeu 180
QY 418 TCCCCCGGGAAATCCCAAGTACCTTGGAGAAGATTAAACAGACATCT 465
Db 181 CysProGlyLysAlaSerSerSerGluLysIleHisGluArgSer 196
RESULT 5
Q9Y444 PRELIMINARY; PRT; 98 AA.
AC Q9Y444;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SYT-SSX protein (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
OK NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=95038836; PubMed=7951320;
RX Clark J., Rocques P.J., Crew A.J., Gill S., Shipley J., Chan A.M.,
RA Gusterson B.A., Cooper C.S.;
RT "Identification of a novel gene, SYT and SSX, involved in the t(X;18)
RT (p11.2;q11.2) translocation found in human synovial sarcoma.";
RL Nat. Genet. 7:502-508(1994).
DR EMBL; X79200; CAB36970.1; -.
FT NON TER 1
SQ SEQUENCE 98 AA; 11136 MW; 1C71F5C8D54513DD CRC64;

Alignment Scores:
Pred. No.: 4,56e-29 Length: 98
Score: 363.00 Matches: 65
Percent Similarity: 91.36% Conservative: 9
Best Local Similarity: 80.25% Mismatches: 7
Query Match: 35.41% Indels: 0
DB: 4 Gaps: 0

US-09-975-856-1 (1-576) x Q9Y444 (1-98)
QY 322 TTCGCGAAGATCATGCCAGAGAGCCAGAGAGAGAAATGTTTGAAGAGAGTCCCA 381
Db :: ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Dy 18 TyrAspGlnIleMetProLysLysProAlaGluGlyAsnAspSerGluGluValPro 37
QY 382 GAGCATCTGCGCCACAAAATGATGGGAAACAGCTGTGCCCGGGGAAATCCAAGTACC 441
Dy ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Dy 38 GluIleSerGlyProGlnAsnAspGlyGlyGluLeuGlySerProGlyLysProThrThr 57
QY 442 TTGAGAGAGATTAACAAGACATCTGGACCCCAAGGGGAAACATGCTCGACCCACAGA 501
Dy ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Dy 58 SerGluLysIleHisGluArgSerGlyProLysArgGlyGluHisAlaTyrThrHisArg 77
QY 502 CTGCGTGAAGAGAGAGCTGCTGTTTATGAGAGATCAGCGACCTGAGGAGATGAC 561
Dy ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Dy 78 LeuArgGluArgLysGlnLeuValIleTyrGluIleLeuSerAspProGluGluAspAsp 97
QY 562 GAG 564
Dy ::::
Dy 98 Glu 98

RESULT 6
Q8WVZ9 PRELIMINARY; PRT; 64 AA.
AC Q8WVZ9;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE BA564H18.1 (Synovial sarcoma, X breakpoint 2) (Fragment).
GN SSX2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Whitehead S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL596242; CAD18884.1; -.
FT NON TER 1
FT NON TER 64
SQ SEQUENCE 64 AA; 7014 MW; 43ADB72AF2FE9613 CRC64;

Alignment Scores:
Pred. No.: 7,93e-18 Length: 64
Score: 255.00 Matches: 64
Percent Similarity: 84.38% Conservative: 48
Best Local Similarity: 75.00% Mismatches: 10
Query Match: 24.88% Indels: 0
DB: 4 Gaps: 4

US-09-975-856-1 (1-576) x Q9NZK4 (1-113)
QY 331 ATCATGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCT 390
Dy ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Dy 1 IleMetProLysLysProAlaGluGlyAsnAspSerGluGluValProGluAlaSer 20
QY 391 GCGCCACAAAATGATGGGAAACAGCTGTGCCCGGGGAAATCCAAGTACCTGGAGAG 450
Dy ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Dy 21 GlyProGlnAsnAspGlyGlyGluLeuGlySerProGlyLysProThrThrSerGluLys 40
QY 451 ATTAACAAGACATCTGGA----- 468
Dy ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Dy 41 IleHisGluArgSerGlyAsnArgGluAlaGlnGluLysGluArgGlyThrAla 60
QY 468 ----- 468
Dy 61 HisArgTyrSerSerGlnAsnThrHisAsnIleGlyArgPheSerLeuSerThrSerMet 80
QY 469 -----CCCAAAAGGGGAAA 483
Dy 81 GlyAlaValHisGlyThrProLysThrIleThrHisAsnArgAspPro-LysGlyGlyAs 100
```

```
Qy 484 CATGCTGGAGCCACAGACTGCGTGAGAGAAAGCAGCTGG 523
|||||
Db 100 nMetProGlyProThrAspCysValArgGluAenSerTrp 113
|||||

RESULT 8
Q80ZT4 PRELIMINARY; PRT; 128 AA.
AC Q80ZT4;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Similar to RIKEN cDNA 4930414C09 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC048441; A48441.1;
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003655; KRAB_related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PS08005; KRAB; 1.
DR PROSITE; PS08006; KRAB_RELATED; 1.
SQ SEQUENCE 128 AA; 15290 MW; F88614D1CBFF3B70 CRC64;

Alignment Scores:
Pred. No.: 1.5e-13 Length: 128
Score: 214.50 Matches: 56
Percent Similarity: 45.00% Conservative: 25
Best Local Similarity: 31.11% Mismatches: 36
Query Match: 20.93% Indels: 63
DB: 11 Gaps: 5

US-09-975-856-1 (1-576) x Q80ZT4 (1-128)
Qy 25 AGGAGACCCAGGATGATGCTCAAAATATCAGAGAAGTTACGAAAGGCTTCGATGATATT 84
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 9 LysAsnProMetGluValLeuTyrGluProLysAsnLeuCysLysAlaPheGlnAspIle 28
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 85 GCCAAATACTTCTTAAGAAAGAGTGGGAAAGATGAATCTCCGGAGAAAATCGTCTAT 144
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 29 SerThrTyrPheSerAspGluGluTrpGlyLysLeuThrGlnLysSerAlaTyr 48
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 145 GTGTATATGAAGTAAACTAGGTCTAGTAACTAGTCTTCAAGGTCCACCTCCCA 204
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 49 ValTyrMetLysArgAsnTyrIleArgMetThrAspLeuGlyValThrValAsnGlnPro 69
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 205 CCTTTCATGCGTAGTAAACCGGCTGCAGACTTCCACGGGAATGATTTGGTAAACGATCGA 264
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 69 ValPheMetArg-----GlyLysGluGln 76
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 265 AACACAGGAATCAGGTTGACGTCCTCAGATGACTTTCGACGCTCCAGAAATCTTC 324
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 77 AspLysGlnSerLeuValGlu----- 83
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 325 CCGAAGATCATGCCCAAGAACCCAGACGAGAGAAATGTTTGAAGGAAGTCCAGAG 384
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 84 -----GlyIle-----GluValHisAsp 89
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 385 GCATCTGGCCCAAAAATGATGGGAAACAGCTGTGCCCCCGGGAAATCCAAAGTACCTTG 444
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 90 Ser----- 90
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Qy 445 GAGAAGATTAAACAGATCTGGACCCCAAGGGGAAACATGCTGGACCCACAGACTG 504
::: ||| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
```

```

Db 121 -----Glytle---GluValHisAsp 126
QY 385 GCATCTGCGCCACAAAATGATGGGAAACAGCTGTGCCCCCGGGAATCAAGTACCTTC 444
Db 127 Ser-----
QY 445 GAGAAGATTAAACAGACATCTGGACCCCAAAAGGGGAAACATGCTGGACCCACAGACTG 504
Db 128 -----AspGluThrSerGlyleArg-----ValAsnValTrpSerHisArgLeu 142
QY 505 CGTCAGAGAAACAGCTGGTGGTTTATCAAGAGATCAGCCACCTCGAGGAAGATGACGAG 564
Db 143 ArgGluArgLysTyrArgValIleTyrGluGluIleSerAspProGluGluGluAsp 162

RESULT 10
Q9CPU1 PRELIMINARY; PRT; 170 AA.
AC Q9CPU1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 4930414C03RIK protein.
GN 4930414C03RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=1217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Flaischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikolaio I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 403:685-690(2001).
DR EMBL; AK015135; BAB29722.1; -.
DR EMBL; AK006218; BAB24465.1; -.
DR MG; MG1:1915235; 4930414C09RIK.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003676; P:nucleic acid binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003855; KRAB-related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PS00805; KRAB; 1.
DR PROSITE; PS00806; KRAB-RELATED; 1.
SQ SEQUENCE 170 AA; 19536 MW; F63F702BFDF7227 CRC64;

Alignment Scores:
Pred. No.: 3,24e-13 Length: 170
Score: 211.50 Matches: 58
Percent Similarity: 47.34% Conservative: 22
Best Local Similarity: 34.32% Mismatches: 62
Query Match: 20.63% Indels: 27
DB: 11 Gaps: 5

US-09-975-856-1 (1-576) x Q9CPU1 (1-170)
QY 67 AAGGCTTCGATGATATTCGCAATACTCTCTAAGAAAGAGTGGGAAAGATGAATCC 126
Db 23 LysAlaPheGlnAspIleSerThrTyrPheSerAspGluGluTrpGlyLysLeuThrGln 42
QY 127 TCGGAGAAATCGTCTATGTGTATAGAACTAAAGCTATGAGGTTCATGACTAAACTAGGT 186
Db 43 TrpGlnLysSerAlaTyrValTyrMetLysArgAsnTyrIleArgMetThrAspLeuGly 62
QY 187 TTCAAGGTACCCCTCCACCTTTCATGCGTAGTAGTAACGGGGCTGCAGAC-----TTC 237
Db 63 ValThrValAsnGlnProValPheMetArgGlyLysGluGlnAlaLysGlnSerLeuVal 82
QY 238 CACGGGAATGATTTTGGTAAACGATCGAACCACAGGATCATCGGTGACCGTCTCAGATG 297
Db 83 GluGlyIleGluVal-----HisAspSerGluAspGluCysPheGluGly 97
QY 298 ACTTTCGCGAGCCTCCAGAGAATCTTCCGGAAGATCATGCCCAAGAACCCAGAGAGAA 357
Db 98 SerPheGly-----ValThrProIleLysArgMetLysLeu 109
QY 358 GAAATGGTTTGAAGGAAGTCCAGAGCGCATCTGCCCCACAAAATGATGGGAACAGCTG 417
Db 110 ThrSerValThrIleSerPheHisAsnValGluGlySerLeuAlaSerGlyGluAsnAsp 129
QY 418 TGCCCCCGGGGAAATCCAAATACCTTGGAGAAGATTAACAAGACATCTGGACCCCAAAAGG 477
Db 130 Cys-----AsnLeuAlaGluThrGlyGlyIleGln--- 139
QY 478 GGGAAACATGCTGCTGCACCCAGACTGCTGAGAGAAAGACAGCTGCTGTTATGAAGAG 537
Db 140 ---ValAsnValTrpSerHisArgLeuArgGluArgTyrArgValIleTyrSerGlu 158
QY 538 ATCAGCGACCCCTGAGGAAGATGACGAG 564
Db 159 IleSerAspThrGluGluGluAsp 167

RESULT 11
O75101 PRELIMINARY; PRT; 117 AA.
AC O75101;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE SSX-HSTT (Fragment).
GN SSX-HSTT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RA Schenke H., Takeuchi T.;
RT "SSX-HSTT";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB012575; BAA32799.1; -.
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 12611 MW; C9F9ADFCD0AD4542 CRC64;

Alignment Scores:
Pred. No.: 6,39e-08 Length: 117
Score: 160.50 Matches: 34
Percent Similarity: 64.62% Conservative: 8
Best Local Similarity: 52.31% Mismatches: 16
Query Match: 15.68% Indels: 7
DB: 4 Gaps: 2

US-09-975-856-1 (1-576) x O75101 (1-117)
QY 241 GGAATGATTTTGGTAAACGATCGAACCACAGGATCAGGTGAACGTCCT-----CAG 294

```


OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wei Y., Sun M., Wang J., Zhu X., Larsson C., Dwight T., Xie Y.,
RA Nilsson G., Larsson O.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY138491; AAN39533.1; -.
FT NON_TER 1
FT NON_TER 39
SQ SEQUENCE 39 AA; 4336 MW; 8FC179F6C8C7E0D CRC64;

Alignment Scores:
Pred. No.: 1.51e-05 Length: 39
Score: 137.00 Matches: 25
Percent Similarity: 87.88% Conservative: 4
Best Local Similarity: 75.76% Mismatches: 4
Query Match: 13.37% Indels: 0
DB: 4 Gaps: 0

US-09-975-856-1 (1-576) x Q8IZG7 (1-39)

QY	322	TTCCCGAAGATCATGCCCAAGCAGCAGCAGGAGAAATGTTTGAAGGAGTGCCA	381
Db	7	TyrAspGlnIleMetProLysLysProAlaGluGluGlyAsnAspSerGluGluValPro	26
QY	382	GAGCATCTGCCCCACAAATGATGGGAAACAGCTGTGC	420
Db	27	GluAlaSerGlyProGlnAsnAspGlyLysGluLeuCys	39

Search completed: March 31, 2004, 13:59:50
Job time : 50 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2004, 15:16:19 ; Search time 2745 Seconds
(without alignments)

6266.162 Million cell updates/sec

Title: US-09-975-856-1

Perfect score: 576

Sequence: 1 ATGAACGGAGACGACGCTT.....ATGACGAGTAACCTCCCTCG 576

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hic.*

9: gb_estl.*

10: gb_est2.*

11: gb_hic.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	496.6	86.2	979	12	BM806411
2	495	836	13	13	BU194624
3	495	887	13	13	BU161779
4	491	85.2	863	13	BQ222907

RESULT 1

BM806411

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM806411 979 bp mRNA linear EST 05-MAR-2002
AGENCOURT_6542819 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5548810
5', mRNA sequence.

BM806411

BM806411.1 GI:19123234

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC <http://mgi.nci.nih.gov/>

1 (bases 1 to 979)

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTF

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM12258 row: f column: 11

High quality sequence stop: 622.

ALIGNMENTS

5	488.6	84.8	1001	12	BM471266
6	481.6	83.6	781	13	BQ434972
7	481.6	83.6	873	13	BU170242
8	481.6	83.6	888	13	BQ432374
9	475.2	83.2	1135	10	BE535379
10	471.6	81.9	602	14	BE158955
11	470.8	81.7	950	13	BQ231752
12	454.6	78.9	793	12	BE560469
13	447	77.6	922	13	BQ229996
14	423.8	73.6	573	10	BE408883
15	423.4	73.5	867	10	BE891434
16	417.6	72.5	627	14	CD767542
17	408.8	71.0	589	10	BE410950
18	402	69.6	943	13	BQ229064
19	400.8	69.6	948	12	BQ333981
20	394.4	68.5	630	10	BE408978
21	394.4	68.5	660	14	CB156144
22	394.4	68.5	759	10	BE384545
23	394.4	68.5	798	10	BE410794
24	388.6	67.5	866	12	EG478295
25	383.4	66.6	636	10	BE390290
26	383	66.5	678	10	BF211314
27	382.4	66.4	757	12	BQ104299
28	377.6	65.6	1053	13	BQ241117
29	373	64.8	848	10	BE387586
30	371.4	64.5	961	13	BU189304
31	370.8	64.4	662	10	BE391023
32	367.6	63.8	1057	12	EG476029
33	360.2	62.5	655	12	BI811262
34	360	62.5	988	10	BE729870
35	360	62.5	1007	12	BQ396977
36	357.6	62.1	1006	13	BU189362
37	340.6	59.1	533	10	BF184274
38	339.4	58.9	958	13	BQ232177
39	330.6	57.4	623	10	BE896093
40	326	56.6	367	10	BE866727
41	324	56.2	927	13	BU178938
42	322	55.9	486	9	AA312651
43	322	55.9	896	10	BE212335
44	319.8	55.5	484	14	CB143725
45	317	55.0	820	10	BE777476

FEATURES source
1. .979
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5548810"
/issue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 86.2%; Score 496.6; DB 12; Length 979;
Best Local Similarity 91.5%; Pred. No. 1.8e-131;
Matches 526; Conservative 0; Mismatches 49; Indels 0; Caps 0;
QY 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATCTCAAAATATCAGAGAAG 60
DB 65 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATCTCAAAATATCAGAGAAG 124
QY 61 TTACGAAAGCCTTCGATGATATGCAAAATATCTTCTTAAGAAAGAGTGGGAAAAGATG 120
DB 125 ATCCAAAAGGCTTCGATGATATGCAAAATATCTTCTTAAGAAAGAGTGGGAAAAGATG 184
QY 121 AAATCCTCGGAGAAATCGTCTATGCTATATGATGCTAACTATGAGTCACTAA 180
DB 185 AAAGCCTCGGAGAAATCTTCTATGCTATATGAGGAAAGATGAGGCTATGACTAA 244
QY 181 CTAGGTTTCAAGGTCACCCCTCCACACCTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
DB 245 CTAGGTTTCAAGGCCACCCCTCCACACCTTTCATGCTAGTAAACGGGCCGAGACTTCCAG 304
QY 241 GGGAAATGTTTGGTACGATCGAACCACAGAAATCAGGTTGAAAGTCTCTCAGATGACT 300
DB 305 GGGAAATGTTTGGTACGATCGAACCACAGAAATCAGGTTGAAAGTCTCTCAGATGACT 364
QY 301 TTCCGACGCTCCAGAGAAATCTTCCGAGATCATGCCCAAGAGCCAGCAGAGGAGAA 360
DB 365 TTCCGACGCTCCAGAGAAATCTTCCGAGATCATGCCCAAGAGCCAGCAGAGGAGAA 424
QY 361 AATGGTTTGAAGGAGTGCAGAGGATCTGGGCCACAAATGATGGGAAACAGCTGTGC 420
DB 425 AATGATTCGGAGGAGTGCAGAGGATCTGGGCCACAAATGATGGGAAACAGCTGTGC 484
QY 421 CCCCAGGAAATCCAACTACCTCTGAGAGATTAACCAAGACATCTGGACCCAAAGGGG 480
DB 485 CCCCAGGAAATCCAACTACCTCTGAGAGATTAACCAAGACATCTGGACCCAAAGGGG 544
QY 481 AAACATGCTGGACCCACAGATCGGTGAGAGAAAGCAGCTGTGTTATGAAGATC 540
DB 545 GAAATGCTGGACCCACAGATCGGTGAGAGAAACAGCTGTGTTATGAAGATC 604
QY 541 AGCCACCTGAGGAGATGACGAGTAATCTCCCTC 575
DB 605 AGCCACCTGAGGAGATGACGAGTAATCTCCCTC 639

RESULT 2
BU194624
LOCUS BU194624.1
DEFINITION AGENCOURT_7969283 NIH_MGC_72 Homo sapiens cdna clone IMAGE:6165903
5', mRNA sequence.
ACCESSION BU194624
VERSION BU194624.1
KEYWORDS EST. GI:22708608
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 836)

AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LML)
DNA Distribution by: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LML at:
<http://image.llnl.gov>
Plate: L1AM13525 row: f column: 16
High quality sequence stop: 624.

FEATURES source

Location/Qualifiers
1. .836
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6165903"
/issue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 85.9%; Score 495; DB 13; Length 836;
Best Local Similarity 91.3%; Pred. No. 4.9e-131;
Matches 525; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 1 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATCTCAAAATATCAGAGAAG 60
DB 74 ATGAACGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATCTCAAAATATCAGAGAAG 133
QY 61 TTACGAAAGCCTTCGATGATATGCAAAATATCTTCTTAAGAAAGAGTGGGAAAAGATG 120
DB 134 ATCCAAAAGGCTTCGATGATATGCAAAATATCTTCTTAAGAAAGAGTGGGAAAAGATG 193
QY 121 AAATCCTCGGAGAAATCGTCTATGCTATATGAGCTAACTATGAGTCACTCACTAAA 180
DB 194 AAAGCCTCGAGAAATCTTCTATGCTATATGAGGAAAGATGAGGCTATGACTAAA 253
QY 181 CTAGGTTTCAAGGTCACCCCTCCACACCTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
DB 254 CTAGGTTTCAAGGCCACCCCTCCACACCTTTCATGCTAGTAAACGGGCCGAGACTTCCAG 313
QY 241 GGGAAATGTTTGGTAAACGATCGAACCACAGGATCAGGTTGAAAGCTCCTCAGATGACT 300
DB 314 GGGAAATGTTTGGTAAACGATCGAACCACAGGATCAGGTTGAAAGCTCCTCAGATGACT 373
QY 301 TTCCGACGCTCCAGAGAAATCTTCCGAGATCATGCCCAAGAGCCAGCAGAGAGAGAA 360
DB 374 TTCCGACGCTCCAGAGAAATCTTCCGAGATCATGCCCAAGAGCCAGCAGAGAGAGAA 433
QY 361 AATGTTTGAAGGAGTGCAGAGGATCTGGGCCACAAATGATGGGAAACAGCTGTGC 420
DB 434 AATGATTCGGAGGAGTGCAGAGGATCTGGGCCACAAATGATGGGAAACAGCTGTGC 493
QY 421 CCCCAGGAAATCCAACTACCTCTGAGAGATTAACCAAGACATCTGGACCCAAAGGGG 480
DB 494 CCCCAGGAAATCCAACTACCTCTGAGAGATTAACCAAGACATCTGGACCCAAAGGGG 553
QY 481 AAACATGCTGGACCCACAGATCGGTGAGAGAAAGCAGCTGTGTTATGAAGATC 540
DB 554 GAAATGCTGGACCCACAGATCGGTGAGAGAAACAGCTGTGTTATGAAGATC 613
QY 541 AGCCACCTGAGGAGATGACGAGTAATCTCCCTC 575
DB 614 AGCCACCTGAGGAGATGACGAGTAATCTCCCTC 648

```
RESULT 3
BUI61779      887 bp  mRNA  linear  EST 04-SEP-2002
LOCUS      BUI61779
DEFINITION  AGENCOURT_7858401 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6168477
5', mRNA sequence.
ACCESSION   BUI61779
VERSION     BUI61779.1 GI:22675689
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 887)
AUTHORS    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL    NIH-MGC http://mgc.nci.nih.gov/.
COMMENT     National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTF
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LUAM13532 row: a column: 22
            High quality sequence stop: 739.
            Location/Qualifiers
                1..887
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:6168477"
                /tissue_type="melanotic melanoma"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH MGC 72"
                /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
                Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
                Average insert size 2 kb. Library constructed by Life
                Technologies."
FEATURES             source
     source
     1..887
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:6168477"
     /tissue_type="melanotic melanoma"
     /lab_host="DH10B (phage-resistant)"
     /clone_lib="NIH MGC 72"
     /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
     Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
     Average insert size 2 kb. Library constructed by Life
     Technologies."
ORIGIN
Query Match      85.9%; Score 495; DB 13; Length 887;
Best Local Similarity 91.3%; Pred. No. 5, 1e-13;
Matches 525; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 1 ATGAACGGAGACGACGCGCTTGCAGAGGAGACCCAGGAGATGCTCAAAATATCAGAGAAG 60
DB 75 ATGAACGGAGACGACGCGCTTGCAGAGGAGACCCAGGAGATGCTCAAAATATCAGAGAAG 134
QY 61 TTACGAAAGGCGCTTCGATGATATTCGCAAAATACCTTCTTAAGAAAGAGTGGGAAAGATG 120
DB 135 ATCCAAAGGCGCTTCGATGATATTCGCAAAATACCTTCTTAAGAAAGAGTGGGAAAGATG 194
QY 121 AAATCCTCGGAGAAATCGTCTATGTTATATGAAGCTAAACTATGAGGTATGACTATAA 180
DB 195 AAAGCCTCAGAGAAATCTTCTATGTTATATGAAGAAAGATGAGGCTATGACTATAA 254
QY 181 CTAGGTTTCAAGGTCACCTCCACCTTTCATGCGTACTTAACGGGTCGACACTCCAC 240
DB 255 CTAGGTTTCAAGGCCACCTCCACCTTTCATGTTGTAATAACGGGCGCAAGACTCCAG 314
QY 241 GCGAATGATTTTGGTAAACGATCGAAACACACAGGAATCAGGTTGAACGTCCTCAGATGACT 300
DB 315 GCGAATGATTTGTAATAATGACCTTAACCGTGGGAATCAGGTTGAACGTCCTCAGATGACT 374
QY 301 TTCGGCAGCCTCCAGAGATCTTCGGAAGATCATGCCCAAGAACCCAGCAGAGGAAGAA 360
DB 375 TTCGGCAGCCTCCAGGAAATCTCCCGGAAGATCATGCCCAAGAACCCAGCAGAGGAAG 434
QY 361 AATGGTTTGAAGGAAGTGCAGAGGATCTGCGCCACAAAAATGATGGGAAACAGCTGTGC 420
```

```
DB 435 AATGATTCGGAGGAAGTCCAGAAAGCATCTCGGCCCAAAATATGATGGAAAGAGACTGTGC 494
QY 421 CCCCGGGAATCCAAAGTACCTTCGAGAGAAATTAAACAGACATCTGGACCCCAAAAGGGGG 480
DB 495 CCCCGGGAATCCAAAGTACCTTCGAGAGAAATTAAACAGAGATCTGGACCCCAAAAGGGGG 554
QY 481 AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAGAGATC 540
DB 555 GAACATGCTGGACCCACAGACTCGGTGAGAGAAACAGCTGGTGTATGAGAGATC 614
QY 541 AGCGACCTCGAGGAAGATGAGAGTAACCTCCCTC 575
DB 615 AGCGACCTCGAGGAAGATGAGAGTAACCTCCCTC 649

BQ222907      863 bp  mRNA  linear  EST 02-MAY-2002
AGENCOURT_7675894 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6095680
5', mRNA sequence.
ACCESSION   BQ222907
VERSION     BQ222907.1 GI:20404307
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 863)
AUTHORS    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL    NIH-MGC http://mgc.nci.nih.gov/.
COMMENT     National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTF
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LUAM13368 row: h column: 17
            High quality sequence stop: 625.
            Location/Qualifiers
                1..863
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:6095680"
                /tissue_type="melanotic melanoma"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH MGC 72"
                /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
                Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
                Average insert size 2 kb. Library constructed by Life
                Technologies."
FEATURES             source
     source
     1..863
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="IMAGE:6095680"
     /tissue_type="melanotic melanoma"
     /lab_host="DH10B (phage-resistant)"
     /clone_lib="NIH MGC 72"
     /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
     Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
     Average insert size 2 kb. Library constructed by Life
     Technologies."
ORIGIN
Query Match      85.2%; Score 491; DB 13; Length 863;
Best Local Similarity 91.2%; Pred. No. 7e-130;
Matches 521; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 5 ACGGAGCAGCGCTTGCAGAGGAGACCCAGGAGATGCTCAAAATATCAGAGAAGTAC 64
DB 1 ACGGAGCAGCGCTTGCAGAGGAGACCCAGGAGATGCTCAAAATATCAGAGAAGTAC 60
QY 65 GAAAGGCGCTTCGATGATTTGCCAAATACCTTCTTAAGAAAGAGTGGGAAAGATGAAT 124
DB 61 AAGAGGCGCTTCGATGATTTGCCAAATACCTTCTTAAGAAAGAGTGGGAAAGATGAAG 120
QY 125 CTTGGAGAAATCGTCTATGTTATATGAGCTTAACCTATGAGGTATGACTAACTAG 184
DB 121 CCTCAGAGAAATCTTCTATGTTATATGAGAGAAAGATGAGGCTATGACTTAACCTAG 180
```

QY 185 GTTCAAGGTCACTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCACGGGA 244
 DB 181 GTTCAAGGCCACCTTCCACCTTTTCATGCTAGTAAACGGGCTGCAGACTTCCACGGGA 240
 QY 245 ATGATTTGGTAAACATCGAAGCAACAGAGATCAGGTTGAAGTCTCAGATGACTTTTCG 304
 DB 241 ATGATTTGGTAAACATCGAAGCAACAGAGATCAGGTTGAAGTCTCAGATGACTTTTCG 300
 QY 305 GCAGGCTCCAGAGATCTTCCGGAAGATCATGCCCAAGAACCCAGCAGAGGAAGAAATG 364
 DB 301 GCAGGCTCCAGGATCTTCCGGAAGATCATGCCCAAGAACCCAGCAGAGGAAGAAATG 360
 QY 365 GTTGAAGGAAGTCCAGAGGATCTGCCCAACAAATGATGGAAACAGTGTGCCCCC 424
 DB 361 ATTCGAGGAAGTGCAGAGATCTGCCCAACAAATGATGGAAACAGTGTGCCCCC 420
 QY 425 CGGGAATCCCAAGTACCTTGGAGAGATTAACAGAGATCTGGACCCAAAGGGGGAAC 484
 DB 421 CGGGAATCCCAAGTACCTTGGAGAGATTAACAGAGATCTGGACCCAAAGGGGGAAC 480
 QY 485 ATGCTGGAACCCACAGACTGCTGTGAGAGAAAGCAGTGTGTTTATGAAGATCAGCG 544
 DB 481 ATGCTGGAACCCACAGACTGCTGTGAGAGAAAGCAGTGTGTTTATGAAGATCAGCG 540
 QY 545 ACCCTGAGAGATCAAGAGTAACCTCCCTC 575
 DB 541 ACCCTGAGAGATCAAGAGTAACCTCCCTC 571

RESULT 5
 BM471266
 LOCUS BM471266 1001 bp mRNA linear EST 05-PEB-2002
 DEFINITION AGENCOURT_6478381 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5563101
 5', mRNA sequence.

ACCESSION BM471266

VERSION BM471266.1 GI:18520308

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 1001)

TITLE NIH-MGC http://mgi.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTF

cDNA library Preparation: Life Technologies, Inc.

cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM1293 row: i column: 22

High quality sequence start: 258

High quality sequence stop: 436

Location/Qualifiers

FEATURES

source

1. .1001

/organism="Homo sapiens"

/db_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5563101"

/tissue_type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 72"

/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: Not I;

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 2 kb. Library constructed by Life

Technologies."

ORIGIN

Query Match

84.8%; Score 488.6; DB 12; Length 1001;

FEATURES

source

1. .781

Best Local Similarity 90.6%; Pred. No. 3.7e-129;
 Matches 521; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACGACCCCTTTCGAAGGAGACCCAGGATGATGCTCAATATACAGAGAG 60
 DB 78 ATGAACGGAGACGACCCCTTTCGAAGGAGACCCAGGATGATGCTCAATATACAGAGAG 137
 QY 61 TTACGAAAGGCTTCGATGATATTCGCAAAATCTTCTCTAAGAAAGAGTGGGAAAGATG 120
 DB 138 ATCCAAAGGCTTCGATGATATTCGCAAAATCTTCTCTAAGAAAGAGTGGGAAAGATG 197
 QY 121 AATCTCTCGGAAATCTGCTATGTTATGTAAGCTTAACTATGAGTCACTGACTTAA 180
 DB 198 AGAGCCTCAGAGAAATCTTCTATGTTATGTAAGAGAAAGTATGAGGCTATGACTAAA 257
 QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
 DB 258 CTAGGTTTCAAGGCAACCTCCACCTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 317
 QY 241 GGAATGATTTGGTAAATCGATCGAATCAGAGATCAGTTGAAGTCTCAGATGACT 300
 DB 318 GGAATGATTTGGTAAATCGATCGAATCAGTTGAAGTCTCAGATGACT 377
 QY 301 TTCGGCAGCTCCAGAGAAATCTTCCGAAAGATCATGCCCAAGAGCCAGCAGAGGAAGA 360
 DB 378 TTCGGCAGATCCAGGAAATCTCCCGAAGATCATGCCCAAGAGCCAGCAGAGGAAGA 437
 QY 361 AATGTTTGAAGGAAGTCCAGAGGATCTGGCCCAAAAATGATGGGAAACAGCTGTGC 420
 DB 438 AATGTTTGAAGGAAGTCCAGAGGATCTGGCCCAAAAATGATGGGAAACAGCTGTGC 497
 QY 421 CCCCAGGAAATCCCAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 480
 DB 498 CCCCAGGAAATCCCAAGTACCTTGGAGAGATTAACAGACATCTGGACCCAAAGGGGG 557
 QY 481 AAACATGCTGACCCACAGACTGCTGTGAGAGAAAGCAGCTGTGTTTATGAAGAGATC 540
 DB 558 GAACATGCTGACCCACAGACTGCTGTGAGAGAAAGCAGCTGTGTTTATGAAGAGATC 617
 QY 541 AGGACCTTGAGGAAGATGACGAGTAACCTCCCTC 575
 DB 618 AGGACCTTGAGGAAGATGACGAGTAACCTCCCTC 652

RESULT 6

BM434972

LOCUS BM434972 781 bp mRNA linear EST 24-MAY-2002

DEFINITION AGENCOURT_7896851 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6159943

5', mRNA sequence.

ACCESSION BM434972

VERSION BM434972.1 GI:21174048

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTF

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1AM13509 row: n column: 08

High quality sequence stop: 634.

Location/Qualifiers

1. .781

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6159943"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 83.6%; Score 481.6; DB 13; Length 781;
Best Local Similarity 89.8%; Pred. No. 3.4e-127;
Matches 517; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACACGCCCTTGCACAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
DB 77 ATGAACGGAGACACACCCTTGCACAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 136
QY 61 TTACGAAAGGCCCTTCGATGATATTGCCAAATACCTCTCTAAGAAAGAGTGGGAAAGATG 120
DB 137 AGAAGCAAGGCCCTTGCATGATATTGCCACATACCTTCTCTAAGAAAGAGTGGGAAAGATG 196
QY 121 AATCTCTCGAGAAATCGTCTATGCTATATGAAGCTAACTATCAGCTCATGACTAAA 180
DB 197 AATATCTCGAGAAATCAGCTATGCTATATGAAGAACTATTAAGGCCATGACTAAA 256
QY 181 CTAGGTTTCAAGGTCACCCCTCCACCTTTCATCGTAGTAAACGGGCTGCAGACTTCCAC 240
DB 257 CTAGGTTTCAAGTCAACCTCCACCTTTCATGCTGTAATAAACAGGCCACAGCTTCCAG 316
QY 241 GGAATGATTTGGTAACGATCGAACCACAGAAATCAGGTTCAAGCTCTCAGATCAT 300
DB 317 GGAATGATTTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 376
QY 301 TTGGGAGCTCCAGAGAACTTTCCGAGATCATGCCAAGAGCCAGCAGAGGAGAA 360
DB 377 TTGGGAGCTCCAGAGATCATCCGAGATCATGCCAAGAGCCAGCAGAGGAGAA 436
QY 361 AATGTTTGAAGAGAGTGCAGAGGATCTGGCCCAAAATGATGGGAAAACAGCTGTGC 420
DB 437 AATGATTCGAAGGAGTGTGAGAAGCATCTGGCCCAAAACGATGGGAAAACAACTGTCAC 496
QY 421 CCCCCGGAATCCAGTCTCTGGAGAGATTAACAAGACATCTGCAGCCCAAAAGGGG 480
DB 497 CCCCCGGAAGCAATATTTCTGAGAAGATTAATAAGAGATCTGACCCCAAAAGGGG 556
QY 481 AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAGAGATC 540
DB 557 AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAGAGATC 616
QY 541 AGCAGCCTGAGGAGATGACGAGTAACTCCCTCG 576
DB 617 AGCAGCCTGAGGAGATGACGAGTAACTCCCTCG 652

RESULT 7
LOCUS BUI70242 873 bp mRNA linear EST 04-SEP-2002
DEFINITION AGENCOURT_7844770 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6051106
5', mRNA sequence.
ACCESSION BUI70242
VERSION BUI70242.1 GI:22684226
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 873)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
DNA distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13304 row: 9 column: 11
High quality sequence stop: 696.

FEATURES
source

1..873
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6051106"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 83.6%; Score 481.6; DB 13; Length 873;
Best Local Similarity 89.8%; Pred. No. 3.6e-127;
Matches 517; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 1 ATGAACGGAGACACGCCCTTGCACAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
DB 80 ATGAACGGAGACACACCCTTGCACAGGAGACCCAGGATGATGCTCAATATCAGAGAAG 139
QY 61 TTACGAAAGGCCCTTCGATGATATTGCCAAATACCTCTCTAAGAAAGAGTGGGAAAGATG 120
DB 140 AGAAGCAAGGCCCTTGCATGATATTGCCACATACCTTCTCTAAGAAAGAGTGGGAAAGATG 199
QY 121 AATCTCTCGAGAAATCGTCTATGCTATATGAAGCTAACTATGAGTCTCATGACTAAA 180
DB 200 AATATCTCGAGAAATCAGCTATGCTATATGAAGAACTATAAGGCCATGACTAAA 259
QY 181 CTAGGTTTCAAGGTCACCCCTCCACCTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
DB 260 CTAGGTTTCAAGTCAACCTCCACCTTTCATGCTGTAATAAACAGGCCACAGCTTCCAG 319
QY 241 GGAATGATTTGGTAACGATCGAACCACAGGATCAGGTTGAAACGTCCTCAGATGACT 300
DB 320 GGAATGATTTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 379
QY 301 TTGGGAGCTCCAGAGAACTTTCCGAGATCATGCCAAGAGTTCGCCAAGAGCCAGCAGAGAGAA 360
DB 380 TTGGGAGGCTCCACAGAAATCATCCGAGATCATGCCAAGAGTTCGCCAAGAGCCAGCAGAGAGAA 439
QY 361 AATGTTTGAAGAGTGCAGAGGATCTGGCCCAAAATGATGGGAAAACAGCTGTGC 420
DB 440 AATGATTCGAAGGAGTGTGAGAAGCATCTGGCCCAAAACGATGGGAAAACAGCTGTGC 499
QY 421 CCCCCGGAATCCAGTCTCTGGAGAAAGTAAACAAGACATCTGGACCCCAAAAGGGG 480
DB 500 CCCCCGAGAAAGCAATATTTCTGAGAAGATTAATAAGAGATCTGGACCCCAAAAGGGG 559
QY 481 AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAGAGATC 540
DB 560 AAACATGCTGGACCCACAGACTCGGTGAGAGAAAGCAGCTGGTGTATGAGAGATC 619
QY 541 AGCAGCCTGAGGAGATGACGAGTAACTCCCTCG 576
DB 620 AGCAGCCTGAGGAGATGACGAGTAACTCCCTCG 655

RESULT 8

B0432374 888 bp mRNA linear EST 24-MAY-2002
 LOCUS AGENCOURT_7859139 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6167925
 5', mRNA sequence.
 ACCESSION B0432374
 VERSION B0432374.1 GI:21171450
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 888)
 NIH-MGC http://mgc.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Straubeberg, Ph.D.
 COMMENT Email: csapob-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHAM13530 row: j column: 22
 High quality sequence stop: 636.
 Location/Qualifiers
 1..888
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6167925"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 72"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

ORIGIN

Query Match	83.6%;	Score 481.6;	DB 13;	Length 888;
Best Local Similarity	89.8%;	Pred. No. 3.6e-127;		
Matches 517;	Conservative 0;	Mismatches 59;	Indels 0;	Gaps 0;

QY 1 ATGAACGGAGACGACGCCCTTTCGAAGGAGACCCAGGATGATGCTCAATATCAGAGAG 60
 DB 79 ATGAACGGAGACGACGCCCTTTCGAAGGAGACCCAGGATGATGCTCAATATCAGAGAG 138
 QY 61 TTACGAAGAGCCCTTCGATGATATGCCAAATACTTCTCTAAGAAAAGATGGGAAAAGATG 120
 DB 139 AGAACCAAGCCCTTTGATGATATTGCCACATACTTCTCTAAGAAAGATGGGAAAAGATG 198
 QY 121 AAATCCTCGAGAAATCGCTATGTGTATATGAAGCTAACTATCAGTCTAGCTATAA 180
 DB 199 AAATACTCGAGAAATCAGCTATGTGTATATGAAGAACTATAAGGCCATGACTATAA 258
 QY 181 CTAGGTTTCAAGGTCAACCTCCACCTTTCATCGGTAGTAAACGGGCTCGACATTTCCAC 240
 DB 259 CTAGGTTTCAAGGTCAACCTCCACCTTTCATGTGTATATGAAGGCCACAGCTTCGAG 318
 QY 241 GCGAATGATTTTGGTAAAGATCGAAACCAAGATCAGGTGTAACGTCCTCAGATGACT 300
 DB 319 GCGAATGATTTTGGTAAATGACCATACCCAGGATTCAGGTTTGAACATCCTCAGATGACT 378
 QY 301 TTCGCGAGCCTCCAGAGATCTTCCCGAGATCATGCCCAAGAGCCACAGAGGAAGAA 360
 DB 379 TTCGCGAGCCTCCAGAGATCATGCCCAAGATCATGCCCAAGAGCCACAGAGGACGAA 438
 QY 361 AATGTTTGAAGGAAGTGCAGAGGATCTGGCCCAAAATATCGGAAAACAGTGTGTC 420
 DB 439 AATGATTGCAAGGAGTGTGCAAGCATCTGGCCCAAAACGATGGGAAACAACTGTGCAC 498

Db 242 CTAGGTTTCAAAAGTCACCTCCACCTTTTCATGTGTATTAATAACAGGCCACAGACTTCCAG 301
 QY 241 GGGATGATTTTGTAAAGATCGAAACACACAGGATCAGGTTTGAACGTCCTCAGATGACT 300
 Db 302 GGAATGATTTTGTAAATGACATTAACCGAGGATTCAGGTTGAACATCTTCATGACT 361
 QY 301 TTGCGCAGCTCCACAGAAATCTTCCGAAGATCATGCCCCAAGAGCCACAGCAGAGAGAA 360
 Db 362 TTGCGCAGCTCCACAGAAATCATCCGAAGATCATGCCCCAAGAGCCACAGCAGAGAGAA 421
 QY 361 AATGTTTGAAGAGAGTCCAGAGCATCTGCCCCACAAATGATGGGAACAGCTGTC 420
 Db 422 AATGATTCGAAGGAGTGCAGAGCATCTGCCCCACAAATGATGGGAACAGCTGTC 481
 QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGGACCCCAAGAGGGG 480
 Db 482 CCCCCAGGAAAGCAATATTTCTGAGAAGATTAATAGAGATCTGGACCCCAAGAGGGG 541
 QY 481 AAACATGCTGACCCACAGATCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 540
 Db 542 AAACATGCTGACCCACAGATCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 601
 QY 541 AGCGACCTGAGGAGATCAGAGTAACCTCCC 572
 Db 602 AGCGACCTGAGGAGATCAGAGTAACCTCCC 633

RESULT 10
 CB158955
 LOCUS
 DEFINITION CB158955 602 bp mRNA linear EST 30-JAN-2003
 K-EST0218399 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-32-F07
 5', mRNA sequence.
 CB158955
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 602)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 21C Frontier Korean EST Project 2001
 Unpublished (2002)
 Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 32 row: F column: 07
 High quality sequence stop: 602.
 Location/Qualifiers
 1. 602
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="L18POOL1n1-32-F07"
 /cell_line="GNU-354-Cho-CK+HLK-3"
 /lab_host="Top10P"
 /clone_lib="L18POOL1n1"
 /note="Organ: Liver; Vector: pT73-Pac; Site: 1: EcoRI;
 Site 2: NotI; The library was contributed by the Soares
 Laboratory and it was constructed as described by Bonaldo,
 M.P., Lennon, G. and Soares, M.B. (1996). Genome Research
 6(9): 791-806. RNA was prepared from harvested cell
 culture."

FEATURES

source
 1. 602
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="L18POOL1n1-32-F07"
 /cell_line="GNU-354-Cho-CK+HLK-3"
 /lab_host="Top10P"
 /clone_lib="L18POOL1n1"
 /note="Organ: Liver; Vector: pT73-Pac; Site: 1: EcoRI;
 Site 2: NotI; The library was contributed by the Soares
 Laboratory and it was constructed as described by Bonaldo,
 M.P., Lennon, G. and Soares, M.B. (1996). Genome Research
 6(9): 791-806. RNA was prepared from harvested cell
 culture."

ORIGIN

Query Match 81.9%; Score 471.6; DB 14; Length 602;
 Best Local Similarity 91.1%; Pred. No. 2.2e-124;

Matches 501; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 QY 1 ATGAACGGAGACACCGCTTTTGAAGGAGACCCAGGATGATGCTCAAATATCAGAGAAG 60
 Db 53 ATGAACGGAGACACCGCTTTTGAAGGAGACCCAGGATGATGCTCAAATATCAGAGAAG 112
 QY 61 TTACGAAAGCCCTTCATGATATTGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 120
 Db 113 ATCAAAAGGCCCTTCATGATATTGCAATATCTCTTAAGAAAGAGTGGGAAAGATG 172
 QY 121 AAATCTCCGAGAGAAATCGTCTATGTGTATATGAAGCTAAACTATAGAGTCACTGACTAA 180
 Db 173 AAAGCTTCGGAGAAATCTTCTATGTGTATATGAAGAGAAAGTATGAGGCTATGACTAA 232
 QY 181 CTAGGTTTCAAGTCAACCTCCACCTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 240
 Db 233 CTAGGTTTCAAGGCCACCTCCACCTTTCATGCTAGTAAACGGGCTGCAGACTTCCAC 292
 QY 241 GGAATGATTTTGGTAAACGATCGAAACACAGGAATCAGGTTGAACGCTCTCAGATGACT 300
 Db 293 GGAATGATTTGATTAATGACCTTAACCGTGGGAATCAGGTTGAACGCTCTCAGATGACT 352
 QY 301 TTGCGCAGCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGAGAGAA 360
 Db 353 TTGCGCAGCTCCAGAGATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGAGAGAA 412
 QY 361 AATGTTTGAAGAGTCCAGAGGATCTGCGCCACAAATCATGAGAAACAGCTGTGC 420
 Db 413 AATGATTCGGAGAGTCCAGAGATCTGCGCCACAAATCATGAGAAAGAGCTGTGC 472
 QY 421 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGAGACCCAAAGAGGGG 480
 Db 473 CCCCAGGAAATCCAAAGTACCTTGGAGAGATTAACAGACATCTGAGACCCAAAGAGGGG 532
 QY 481 AAACATGCTGACCCACAGATCTGCGCCACAAATCATGAGAAACAGCTGTGC 540
 Db 533 GAACATGCTGACCCACAGATCTGCGCCACAAATCATGAGAAACAGCTGTGC 592
 QY 541 AGCGACCTG 550
 Db 593 AGCGACCTG 602

RESULT 11

BQ231752
 LOCUS
 DEFINITION BQ231752 950 bp mRNA linear EST 02-MAY-2002
 5', mRNA sequence.
 ACCESSION BQ231752
 VERSION BQ231752.1 GI:20413152
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 950)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1343 row: i column: 08
 High quality sequence stop: 532.
 Location/Qualifiers
 1. 950
 /organism="Homo sapiens"

REFERENCE

AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES

source


```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:606127"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN
Query Match      81.7%; Score 470.8; DB 13; Length 950;
Best Local Similarity 89.9%; Pred. No. 4.8e-124;
Matches 516; Conservative 0; Mismatches 57; Indels 1; Gaps 1;

QY 1 ATGAACGGAGACGACGCTTTGCAAGAGACCCAGGATGATGCTCAATATCAGAGAAG 60
DB 74 ATGAACGGAGACGACGCTTTGCAAGAGACCCAGGATGATGCTCAATATCAGAGAAG 133
QY 61 TTACGAAAGCCTTCGATGATATTGCCAATACTTCTTAAGAAAGATGGGAAAGATG 120
DB 134 ATCAAAAGGCTTCGATGATATTGCCAATACTTCTTAAGAAAGATGGGAAAGATG 193
QY 121 AATCTCTCGAGAAATCGTCTATGTATATGAAGCTAACTATGAGCTCATGACTAA 180
DB 194 AAGCCTCAGAGAAATCTTCTATGTATATGAAGAAAGATGATGAGCTATGACTAA 253
QY 181 CTAGGTTTCAAGCTCACCCTCCACCTTTTCATCGTAGTAAGAAAGGCTGCAGACTTCCAC 240
DB 254 CTAGGTTTCAAGGCTCACCCTCCACCTTTTCATCGTAGTAAGAAAGGCTGCAGACTTCCAC 313
QY 241 GGAATGATTTTGGTAGCATGAAACCCAGAGAAATCAGGTGAAGCTCCTCAGATGACT 300
DB 314 GGAATGATTTTGGTAGCATGAAACCCAGAGAAATCAGGTGAAGCTCCTCAGATGACT 373
QY 301 TTGGCAGCTCCAGAGAAATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 360
DB 374 TTGGCAGCTCCAGAGAAATCTTCCGAGAGATCATGCCCAAGAGCCAGCAGAGGAAGAA 433
QY 361 AATGTTTGAAGAGTGCAGAGGATCTGGGCCCAAAATGATGGGAAACAGCTGTGC 420
DB 434 AATGATTTCGAGGAAGTGCAGAGGATCTGGGCCCAAAATGATGGGAAACAGCTGTGC 493
QY 421 CCCCCGGAAATCAAGTACCTTGGAGAGATTAAAGACATCTGGACCCCAAAAGGGGG 480
DB 494 CCCCCGGAAACCAACTACTCTGAGAGATTTCAGAGAGATCTGGACCCCAAAAGGGGG 553
QY 481 AAACAT-GCCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATGAGAGAT 539
DB 554 GAACATGCTGGACCCACAGACTGCGTGAGAGAAACAGCTGGTGTATGAGAGAT 613
QY 540 CAGGACCTCGAGGAAGATGACGAGTAACCTCCC 573
DB 614 CAGGACCTCGAGGAAGATGACGAGTAACCTCCC 647

RESULT 12
BI560469
LOCUS 603254435F1 NIH_MGC_97 Homo sapiens CDNA clone IMAGE:5296880 5',
DEFINITION mRNA sequence.
ACCESSION BI560469
VERSION BI560469.1 GI:1544770
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 793)
NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
```

Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11751 row: e column: 09
High quality sequence stop: 790.

FEATURES
Location/Qualifiers
1..793
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5296880"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"
/notes="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
(gtcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.2 kb and normalized to R07 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN
Query Match 78.9%; Score 454.6; DB 12; Length 793;
Best Local Similarity 90.1%; Pred. No. 2e-119;
Matches 520; Conservative 0; Mismatches 54; Indels 3; Gaps 3;
QY 1 ATGAACGGAGACGACGCTTTGCAA-GGAGACCCAGGATGATGCTCAATATCAGAGAA 59
DB 143 ATGAACGGAGACGACGCTTTGCAAGGGAGACCCACGTTGGTCTCAAAATACAGAGAA 202
QY 60 GTTACGAAAGCCTTCGATGATATTGCCAATACTTCTTAAGAAAGATGGGAAAGAT 119
DB 203 GATCCAAAGGCTTCGATGATATTGCCAATACTTCTTAAGAAAGATGGGAAAGAT 262
QY 120 GAAATCTCTCGAGAAATCGTCTATGTATATGAAGCTAAACTATGAGTCTGACTAA 179
DB 263 GAAGCCTCGAGAAATCTTCTATGTATATGAAGAAAGTATGAGGCTATGACTAA 322
QY 180 ACTAGTTTCAAGTCAACCTCCACCTTTTCATCGTAGTAAGACGGCTGCAGACTTCCA 239
DB 323 ACTAGTTTCAAGGACACCTCCACCTTTTCATGTGTAATAACGGCCGAGAGACTTCCA 382
QY 240 CCGGAATGATTTTGTATACGATCGAAACCCAGAGAAATCAGGTTGAACGTCCTCAGATGAC 299
DB 383 -GGGAATGATTTGGATATGATACCTAACCGTGGGAATCAGGTTGAACGTCCTCAGATGAC 441
QY 300 TTTCGGAGCTTCAGAGAAATCTTCCGAGATCATGCCCAAGAGCCAGCAGAGGAGA 359
DB 442 TTTCGGAGCTTCAGAGAAATCTTCCGAGATCATGCCCAAGAGCCAGCAGAGGAGA 501
QY 360 AAATGTTTGAAGGAAGTGCAGAGGATCTTGGCCCAAAATGATGGGAAACAGCTGTG 419
DB 502 AAATGATTCGGAGGAAGTGCAGAGGATCTTGGCCCAAAATGATGGGAAAGAGCTGTG 561
QY 420 CCCCCGGGAAATCAAGTACTCTTGGAGAAAGATTAAAGACATCTGACCCCAAGGGG 479
DB 562 CCCCCGGGAAACCAACTACTCTGAGAGATTTACAGAGAGATCTGACCCCAAGGGG 621
QY 480 GAACATGCTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGTATGAA-GAGA 538
DB 622 GAACATGCTGGACCCACAGACTGCGTGAGAGAAACAGCTGGTGTATGAAACGAGA 681
QY 539 TCAGCAGCCTCGAGGAAGATGACGAGTAACCTCCCCTC 575


```
|||||
692 TCAGCGACCTGAGAAAGATGACGAGTAACCTCCCTC 718

RESULT 13
BQ229996      922 bp      mRNA      linear      EST 02-MAY-2002
LOCUS      AGENCOURT_7560406 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6048424
DEFINITION      5', mRNA sequence.
ACCESSION      BQ229996
VERSION      BQ229996.1 GI:20411396
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      NIH-MGC http://mgi.nci.nih.gov/
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE      Unpublished (1999)
JOURNAL      Contact: Robert Strausberg, Ph.D.
COMMENT      Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13297 row: g column: 17
High quality sequence stop: 483.

FEATURES
Source
1..922
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6048424"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/notes="Organ: Skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN
Query Match      77.6%; Score 447; DB 13; Length 922;
Best Local Similarity 89.1%; Pred. No. 3.2e-117;
Matches 516; Conservative 0; Mismatches 60; Indels 3; Gaps 3;

1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 60
44 ATGAACGGAGACGACACCTTTGCAAGGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 103
61 TTAGGAAGCCCTTCGATGATATGCAATATCTCTTAAGAAAGAGTGGGAAGATG 120
104 AGAAGCAAGCCCTTCGATGATATGCAATATCTCTTAAGAAAGAGTGGGAAGATG 163
121 AATATCTCGGAGAAATCGCTATGTATATGAAGCTAAATATGAGTCTAGCTAA 180
164 AATATCTCGGAGAAATCAGCTATGTATATGAAGAAATATGAAGCCATGACTAAA 223
181 CTAGGTTCAAGTCAACCTCCACCTTTATCGGTAGTAAACGGGTGACAGCTCCAC 240
224 CTAGGTTCAAGTCAACCTCCACCTTTATCGGTAGTAAACGGGTGACAGCTCCAC 283
241 GGGATGATTTTGTATGATGACCAATACCGGAGATTGAGTTCCTTCAGATGACT 300
284 GGGATGATTTTGTATGATGACCAATACCGGAGATTGAGTTCCTTCAGATGACT 343
301 TTCCGCGACCTCCAGAGAAATCTTCCGAGATCATGCCCCAAGAGCCAGAGGAGAA 360
344 TTCCGCGACCTCCAGAGAAATCTTCCGAGATCATGCCCCAAGAGCCAGAGGAGAA 403
```

```
|||||
361 AATGTTTGAAGAGTCCAGAGGATCTGCCCCACAAATATGTTGGGAAACAGCTGTGC 420
404 AATGATTCGAAGGGAGTGTCAAGAGCATCTGCCCCACAAACGATGGGAAACACTGCAC 463
421 CCCCCGGGAAATCCAAAGTACCTTGGGAGAGATTAAACAGACATCTGGACCCCAAAA- 479
464 CCCCAGGAAAGCAATTAATCTTGAGAGATTAAATAGAGATCTGACCCCAAGGGGG 523
480 GAAACATGCTGACCCACAGACTGCTGAGAGAAAGCAGCTGGTGG-TTTATGAGAGA 538
524 GAAACATGCTGACCCACAGACTGCTGAGAGAAAGCAGCTGGGGAATTTATGAGAGA 583
539 TCAG-CGACCTCAGGAAGATGACGAGTAACCTCCCTCG 576
584 TCAGCCGACCTCAGGAAGATGACGAGTAACCTCCCTCG 622

RESULT 14
BE408883      573 bp      mRNA      linear      EST 21-JUL-2000
LOCUS      601303758F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3637982 5',
DEFINITION      mRNA sequence.
ACCESSION      BE408883
VERSION      BE408883.1 GI:9345333
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      NIH-MGC http://mgi.nci.nih.gov/
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE      Unpublished (1999)
JOURNAL      Contact: Robert Strausberg, Ph.D.
COMMENT      Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM338 row: d column: 15
High quality sequence stop: 571.

FEATURES
Source
1..573
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3637982"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 21"
/notes="Organ: Placenta; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN
Query Match      73.6%; Score 423.8; DB 10; Length 573;
Best Local Similarity 88.9%; Pred. No. 1.2e-110;
Matches 458; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 60
59 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGAGTATGCTCAAAATATCAGAGAAG 118
61 TTACGAAGCCCTTCGATGATATGCAATATCTCTTAAGAAAGAGTGGGAAGATG 120
```

Db 119 AGAAGCAAGGCTTTGATGATATTGCCACATCTTCTTAAGAAAGAGTGGAAAAAGATG 178
QY 121 AATCTCTCGAGAAATCGCTATGTATATGATGCTAACTATGAGTCTATGACTTAA 180
Db 179 AATATCTCGAGAAATCAGCTATGTATATGAGAGAACTATAGGCCATGACTTAA 238
QY 181 CTAGGTTTCAAGTCACTCCCTCCACCTTTATCGGTAGTAAACGGGCTCCAGACTTCCAC 240
Db 239 CTAGGTTTCAAGTCACTCCCTCCACCTTTATCGGTAGTAAACGGGCTCCAGACTTCCAG 298
QY 241 GGAATGATTTTGGTACGATCGAACCACAGCAATCAGGTTGAACGTCCTCAGATGACT 300
Db 299 GGAATGATTTTGAATGATACCAATACCGCAGATTCAGGTTGAACATCTCAGATGACT 358
QY 301 TTCGGCAGCCTCCAGAGAAATCTTCCGGAAGATCATGCCCAAGAGCCAGCAGAGGAGAA 360
Db 359 TTCGGCAGGCTCCACAGATCATCTCCGGAAGATCATGCCCAAGAGCCAGCAGAGGAGAA 418
QY 361 AATGGTTTGAAGAGTCCAGAGGATCTGGCCCAACAAATGATGGGAAACAGCTGTGC 420
Db 419 AATGATTCGAAGGAGTGTGAGAGCATCTGGCCCAACAAATGATGGGAAACAGCTGTGC 478
QY 421 CCCCCGGGAAATCCAAAGTCTTGGAGAGATTAAACAAGACATCTTGGACCCCAAAAGGGGG 480
Db 479 CCCCCGGAAGAAATATTTCTGAGAGATTAAACAAGATCTTGGACCCCAAAAGGGGG 538
QY 481 AAACATGCTTGGACCCACAGACTGCGTGAGAGAA 515
Db 539 AAACATGCTTGGACCCACAGACTGCGTGAGAGAA 573

RESULT 15
BE891434
LOCUS 601434893F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3919991 5',
- DEFINITION mRNA sequence.
ACCESSION BE891434
VERSION BE891434.1 GI:10350764
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 867)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9750 row: j column: 24
High quality sequence stop: 641.
Location/Qualifiers
1..867
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3919991"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo.dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

FEATURES
source

ORIGIN

Query Match 73.5%; Score 423.4; DB 10; Length 867;
Best Local Similarity 89.4%; Pred. No. 1.9e-110;
Matches 512; Conservative 0; Mismatches 56; Indels 5; Gaps 5;
QY 1 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATCTCAATATCAGAGAAG 60
Db 50 ATGAACGGAGACGACGCTTTGCAAGGAGACCCAGGAGATGATCTCAATATCAGAGAAG 109
QY 61 TTACGAAGAGCCTTCGATGATATTGCCAAATACTTCTCTAAGAAAGAGTGGGAAAAAGATG 120
Db 110 ATCCAAAGAGCCTTCGATGATATTGCCAAATACTTCTCTAAGGAGAGTGGGAAAAAGATG 169
QY 121 AATTCCTCGAGAAATCGTCTATGTATATGAAGCTAACTATGAGTCTCAGACTAAA 180
Db 170 AAGCCCTCAGAGAAATCTTCTATGTATATGAAGAAAGATGAGGCTATGACTTAA 229
QY 181 CTAGGTTTCAAGGTCACCCCTCCACCTTTTCATGCTAGTAAACGGGCTCGAGACTTCCAC 240
Db 230 CTAGGTTTCAAGGTCACCCCTCCACCTTTTCATGCTAGTAAACGGGCTCGAGACTTCCAC 288
QY 241 GGAATGATTTTGGTAAAGATCGAAACCCAGGATCAGGTTGACGTCCTCAGATGACT 300
Db 289 GGAATGATTTTGGTAAAGATCGAAACCCAGGATCAGGTTGACGTCCTCAGATGACT 348
QY 301 TTCGGCAGCCTCCAGAGAAATCTTCCGGAAGATCATGCCCAAGAGCCAGCAGAGGAGAA 360
Db 349 TTCGGCAGGCTCCAGGAAATCTTCCGGAAGATCATGCCCAAGAGCCAGCAGAGGAGAA 408
QY 361 AATGGTTTGAAGAGTCCAGAGGATCTGGCCCAACAAATGATGGGAAACAGCTGTGC 420
Db 409 AATGATTCGAAGGAGTCCAGAGATCTTGGCCCAACAAATGATGGGAAAGAGCTGTGC 468
QY 421 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAACAAGACATCTTGGACCC-AAAAAGGGG 479
Db 459 CCCCCGGGAAATCCAAAGTACCTTGGAGAGATTAAACAAGACATCTTGGACCC-AAAAAGGGG 527
QY 480 GAAACATGCTTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGAT 539
Db 528 GGAACATGCTTGGACCCACAGACTGCGTGAGAGAAAGCAGCTGGTGGTTTATGAAGAGAT 586
QY 540 CAGCGACCTCGAGGAGATGACGAGTAACCTCC 572
Db 587 CAGCGA-CCTGAGGAGATGACGAGTAACCTCC 618
Search completed: March 31, 2004, 17:25:05
Job time : 2750 secs